

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE CIENCIAS FÍSICAS**  
**Departamento de Óptica**



**REGISTRO AUTOMÁTICO DE IMÁGENES DE BANDAS  
ESPECTRALES ALEJADAS**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR  
PRESENTADA POR**

Enrique Coiras López

Bajo la dirección del doctor

Javier Santamaría

**Madrid, 2003**

**ISBN: 978-84-669-1568-7**

**©Enrique Coiras López, 1999**



Facultad de Ciencias Físicas

Facultad de Ciencias Físicas  
Departamento de Óptica

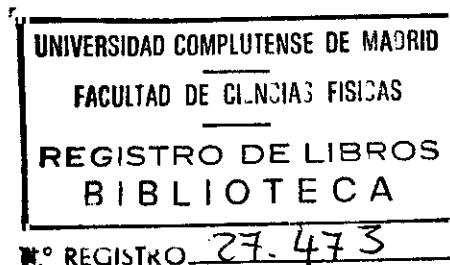
Universidad Complutense  
de Madrid



Universidad  
Complutense



# Registro Automático de Imágenes de Bandas Espectrales Alejadas



Memoria presentada en la Facultad de Ciencias Físicas  
de la Universidad Complutense de Madrid por

**Enrique Coiras López**

para optar al grado de Doctor en Ciencias Físicas

Madrid, Agosto 1999

**Departamento de Optica**  
Facultad de Ciencias Físicas  
Universidad Complutense de Madrid

Tesis Doctoral

**Registro Automático de Imágenes  
de Bandas Espectrales Alejadas**

Realizada en SENER, Ingeniería y Sistemas  
con beca del Ministerio de Educación y Cultura IN92-D50837758

Madrid, Agosto 1999

## Resumen

En esta memoria se presentan nuevos métodos para el registro de imágenes de bandas espectrales alejadas. Los niveles de gris o las texturas de las imágenes no pueden utilizarse para el registro de imágenes de diferente naturaleza, por lo que se han desarrollado dos procedimientos de registro basados en el casamiento de bordes.

Las imágenes de bordes se procesan para extraer segmentos lineales rectos, que son posteriormente agrupados para formar triángulos. El conjunto de transformaciones candidatas para el registro de las imágenes se determina mediante el casamiento de un triángulo de la imagen fuente con otro de la imagen destino. Los criterios de casamiento para segmentos procedentes de imágenes similares (coincidencia de vértices, longitud de superposición, diferencias de contraste, etc.) no son válidos para su aplicación a segmentos procedentes de imágenes de distintas bandas espectrales, por lo que se ha desarrollado un nuevo criterio de casamiento para segmentos que depende únicamente de la probabilidad de que las rectas que definan los segmentos se correspondan con un mismo borde físico de la escena real.

Se estudia también la aplicación al registro de las técnicas de acumulación empleadas en la transformada de Hough. A partir de ellas, se ha desarrollado un procedimiento específico de registro para imágenes visible/IR con imágenes SAR.

Para la evaluación de los procedimientos de registro desarrollados, se describen dos métodos diferentes. El primero se basa en la utilización de la transformada de Gabor para la determinación del incremento de energía, en la banda de frecuencias espaciales más altas, causado por la duplicación de bordes asociada al desregistro. El segundo emplea técnicas de correlación y flujo óptico para la determinación del campo vectorial de desplazamientos locales correspondientes a la deformación de desregistro.

Finalmente se presentan y analizan los resultados de la aplicación de los métodos de registro desarrollados a varios tipos de imágenes (en particular, visible-visible, visible-IR y visible-SAR). También se proponen otras posibles aplicaciones de los métodos de registro desarrollados.

**Palabras clave:** Registro de imágenes, registro basado en bordes, registro basado en segmentos, casamiento de segmentos, registro visible-IR, registro de imágenes SAR.



## Abstract

New general registration methods for images of different nature are presented in this document. As grey-levels or textures cannot be used for the registration of images from separate spectral bands, two edge-based registration methods have been developed.

Edge images are processed to extract straight linear segments, which are then grouped to form triangles. A set of candidate transformations is determined by matching triangles from the source and destination images. The transformations are then evaluated by matching the transformed set of source segments to the set of destination segments. As the coincidence of vertices or edge overlapping cannot be assumed in the registration of images of different nature, a new function for evaluating the matching quality between source and destination segments which does not rely on overlapping measures is proposed.

The application of Hough techniques for the determination of the best global registration transformation is also studied, and used for the development of a new visible/IR-SAR registration method.

For the validation of the registration procedures, two new evaluation methods have been developed. The first is based on measuring the energy increase on the high spatial frequencies associated with missregistration. The second one consists of the determination of the vector field of local displacements corresponding to the misregistration deformation.

Results and objective evaluation of the registration methods for several types of images (in particular, visual-visual, visual-IR and visual-SAR) are presented. Some other applications for the developed registration method are proposed.

**Keywords:** Image registration, edge-based registration, segment-based registration, segment matching, visual-IR registration, SAR registration.

## Agradecimientos

En primer lugar, me gustaría expresar mi más sincero agradecimiento a Javier Santamaría, director de este trabajo, por su constante apoyo e inagotable paciencia. Me considero muy afortunado por haber tenido la oportunidad de trabajar bajo su tutela y espero poder seguir aprendiendo de él mucho tiempo. Este trabajo es, además, parte de la línea de investigación en la que él lleva trabajando desde hace años.

Quisiera ofrecer un agradecimiento especial a Carlos Miravet, por sus inestimables enseñanzas y aportaciones. Es imposible calificar el trabajo descrito en esta tesis como realizado por una única persona y la importante colaboración de Carlos Miravet ha sido fundamental en su desarrollo.

Este trabajo es, por otra parte, una lógica continuación de las investigaciones realizadas previamente por Maite Gómez y Alberto Domingo, asimismo bajo la dirección de Javier Santamaría, en el campo de la Fusión de Imágenes. Quiero también agradecerles a ellos la gran ayuda que me han prestado, principalmente en las etapas iniciales del trabajo.

Todo este proyecto habría sido imposible sin el acuerdo de colaboración entre SENER, Ingeniería y Sistemas, y la Universidad Complutense de Madrid. Por ello, me gustaría agradecer a D. Manuel Valls y a D. Rafael Quintana, responsables del Departamento y la División Aeroespacial en SENER, no sólo el apoyo a la iniciativa de colaboración, sino también el constante interés mostrado por el proyecto mismo. Por parte de la Universidad, debo hacer una mención especial al profesor Eusebio Bernabeu, director del Departamento de Óptica, que, además de aceptar el proyecto, fue quien me inició, en sus clases junto con Luis Sánchez Soto, a esta maravillosa y apasionante rama de la Física que es la Óptica. Asimismo, quiero dar gracias al profesor Jesús Marcén, Tutor de la tesis, por su valiosa orientación y ayuda.

También les debo especial gratitud a Rafael Navarro y a su equipo del Instituto de Óptica (Oscar Nestares, Antonio Tabernero y Javier Portilla, en particular), ya que una parte muy importante de este trabajo se apoya en las investigaciones que han realizado en el campo del análisis de la visión humana. He aprendido mucho, tanto en las conversaciones que he mantenido con ellos, como en el estudio de sus tesis doctorales y artículos.

Hay mucha más gente que me ha ayudado o ha colaborado, de muy variadas e indirectas formas, al desarrollo de esta tesis. Sería imposible nombrarlos a todos, pero me gustaría mencionar especialmente a mis padres, a los que les dedico cariñosamente esta memoria.

# Indice

- 1. INTRODUCCIÓN ..... 1
  - 1.1. JUSTIFICACIÓN Y PROPÓSITO ..... 4
  - 1.2. ESTRUCTURA DE LA MEMORIA..... 5
- 2. REGISTRO DE IMÁGENES..... 7
  - 2.1. PROCEDIMIENTOS GENERALES DE REGISTRO ..... 7
    - 2.1.1. Registro manual..... 7
    - 2.1.2. Registro automático..... 9
  - 2.2. ANTECEDENTES..... 10
    - 2.2.1. Registro de imágenes similares o de bandas espectrales próximas ..... 11
    - 2.2.2. Registro de imágenes de bandas espectrales alejadas ..... 12
- 3. EXTRACCIÓN DE ESTRUCTURAS GEOMÉTRICAS Y DETERMINACIÓN DE TRANSFORMACIONES..... 17
  - 3.1. ELEMENTOS INVARIANTES ..... 17
  - 3.2. PREPROCESADO DE LAS IMÁGENES ..... 20
    - 3.2.1. Filtrado de ruido..... 20
    - 3.2.2. Extracción de bordes y detección de segmentos lineales ..... 23
  - 3.3. TRANSFORMACIONES ASOCIADAS A CASAMIENTOS LOCALES DE ESTRUCTURAS GEOMÉTRICAS ..... 27
    - 3.3.1. Casamiento de triángulos ..... 27
    - 3.3.2. Casamiento de parejas de segmentos lineales..... 33
    - 3.3.3. Casamiento de segmentos lineales aislados ..... 35
  - 3.4. EVALUACIÓN DE LAS TRANSFORMACIONES DE REGISTRO..... 36
- 4. REGISTRO AUTOMÁTICO BASADO EN TÉCNICAS DE ACUMULACIÓN.. 37
  - 4.1. INTRODUCCIÓN. TRANSFORMADA DE HOUGH ..... 37
  - 4.2. TÉCNICAS DE ACUMULACIÓN APLICADAS AL REGISTRO DE IMÁGENES ..... 39
    - 4.2.1. Descripción general del método de acumulación..... 39
    - 4.2.2. Determinación de las precisiones de acumulación ..... 41
  - 4.3. REDUCCIÓN DE LA DIMENSIONALIDAD DEL ESPACIO DE ACUMULACIÓN..... 47
  - 4.4. DETERMINACIÓN DE LA FUNCIÓN DE PESO ..... 50
  - 4.5. ACUMULACIÓN DE TRANSFORMACIONES AFINES INCOMPLETAS..... 54
  - 4.6. VALORACIÓN Y LIMITACIONES DEL MÉTODO DE ACUMULACIÓN ..... 57

<b>5. REGISTRO AUTOMÁTICO BASADO EN EL CASAMIENTO DE CONJUNTOS DE SEGMENTOS LINEALES.....</b>	<b>59</b>
5.1. INTRODUCCIÓN .....	59
5.2. CALIDAD DE CASAMIENTO DE DOS SEGMENTOS LINEALES .....	60
5.2.1. <i>Antecedentes</i> .....	60
5.2.2. <i>Calidad basada en superposición</i> .....	62
5.2.3. <i>Calidad basada en distancia</i> .....	64
5.3. CALIDAD DE CASAMIENTO DE DOS CONJUNTOS DE SEGMENTOS LINEALES .....	70
<b>6. CÁLCULO DE TRANSFORMACIONES DE REGISTRO DE ORDENES SUPERIORES .....</b>	<b>75</b>
6.1. INTRODUCCIÓN .....	75
6.2. REFINAMIENTO DE LA TRANSFORMACIÓN AFÍN ÓPTIMA .....	75
6.2.1. <i>Ejemplos de aplicación del método de refinamiento</i> .....	78
<b>7. EVALUACIÓN GLOBAL DEL REGISTRO .....</b>	<b>81</b>
7.1. EVALUACIÓN BASADA EN FUSIÓN DE IMÁGENES POR DESCOMPOSICIÓN DE GABOR ...	81
7.1.1. <i>Fusión de imágenes basada en ondículas de Gabor</i> .....	83
7.1.2. <i>Criterio de evaluación del registro</i> .....	87
7.1.3. <i>Imágenes de prueba</i> .....	90
7.2. EVALUACIÓN BASADA EN LA DETERMINACIÓN DEL CAMPO VECTORIAL DE DESPLAZAMIENTOS LOCALES .....	101
7.2.1. <i>Obtención de un valor de desplazamiento medio equivalente</i> .....	101
7.3. EJEMPLO DE APLICACIÓN DE LOS CRITERIOS DE EVALUACIÓN DESARROLLADOS.....	109
<b>8. APLICACIONES DEL REGISTRO AUTOMÁTICO .....</b>	<b>115</b>
8.1. REGISTRO VISIBLE-IR .....	115
8.2. REGISTRO VISIBLE/IR-SAR .....	121
8.2.1. <i>Introducción</i> .....	121
8.2.2. <i>Descripción general del método de registro visible/IR-SAR</i> .....	121
8.2.3. <i>Refinamiento del registro mediante acumulación</i> .....	123
8.2.3. <i>Función de peso para la localización de segmentos</i> .....	125
8.3. OTRAS APLICACIONES .....	133
8.3.1. <i>Reconocimiento y detección de formas</i> .....	133
8.3.2. <i>Mosaicos de imágenes</i> .....	134
8.3.3. <i>Registro de mapas y fotografías</i> .....	135
<b>9. CONCLUSIONES.....</b>	<b>137</b>

<b>APÉNDICE A: MODELO DE TRANSFORMACIÓN AFÍN UTILIZADO.....</b>	<b>141</b>
<b>APÉNDICE B: REDUCCIÓN DEL RUIDO DE BARRIDO EN IMÁGENES IR...</b>	<b>143</b>
<b>APÉNDICE C: CRECIMIENTO HEXADECAGONAL DE REGIONES .....</b>	<b>147</b>
C.1. INTRODUCCIÓN.....	147
C.2. CRECIMIENTO OCTOGONAL REGULAR.....	148
C.3. MODELO PARA EL CRECIMIENTO HEXADECAGONAL .....	150
C.4. RESULTADOS Y APLICACIONES .....	153
<b>BIBLIOGRAFÍA .....</b>	<b>157</b>
<b>INDICE DE FIGURAS. ....</b>	<b>163</b>
<b>INDICE DE TABLAS. ....</b>	<b>171</b>
<b>NOTAS .....</b>	<b>173</b>

# 1. Introducción

Los sistemas de adquisición de imágenes se utilizan para obtener información sobre una escena de forma remota. El proceso de obtención de la información se fundamenta básicamente en el análisis de la radiación electromagnética (principalmente) proveniente de la escena, que es recogida por un detector después de haberse propagado a través de un determinado medio de transmisión. El tipo de información obtenida depende de la escena observada, del sistema de observación, del medio de transmisión y de la banda del espectro electromagnético a la que sea sensible el detector.

Históricamente, la banda de longitudes de onda correspondiente al espectro visible ha sido la más utilizada, por ser la banda de funcionamiento del ojo humano y coincidir, además, con el rango de máxima irradiancia solar. Los sistemas que funcionan en el rango visible presentan costes de fabricación y mantenimiento relativamente bajos, su resolución es alta en condiciones atmosféricas buenas, y son poco sensibles a los cambios de temperatura y humedad. Además, las imágenes obtenidas son fácilmente interpretables, por su semejanza con el aspecto visual de la escena. Como inconvenientes pueden citarse su dependencia de las condiciones de iluminación, que limita su capacidad de funcionamiento en situaciones de baja luminosidad, y sus bajas prestaciones en condiciones de baja transmitancia atmosférica, o en presencia de humo, niebla o polvo en suspensión.

Más recientemente se ha extendido la utilización de sensores de formación de imágenes que operan en otras bandas espectrales. Estos sensores proporcionan información adicional a la que se obtiene mediante el análisis de la radiación visible únicamente. Es el caso de los sistemas infrarrojos (IR) en las bandas espectrales de 3 a 5  $\mu\text{m}$  y de 8 a 12  $\mu\text{m}$ , que corresponden a ventanas de transmitancia atmosférica [1], y que cuentan con aplicaciones en medicina, industria y defensa, principalmente. Los sistemas infrarrojos tienen como principal ventaja su independencia de fuentes de iluminación externa, sobre todo en la región de infrarrojo lejano (8-12  $\mu\text{m}$ ), lo que les permite funcionar tanto de día como de noche. Además, son menos sensibles a la presencia de niebla, bruma o polvo en suspensión. Como contrapartida, estos sistemas suelen ser más caros y difíciles de mantener, presentan una resolución espacial menor y las imágenes pueden sufrir degradaciones en condiciones de alta temperatura y humedad.

Si bien tradicionalmente la adquisición de información ha estado basada en la utilización de un único sensor específico, la tendencia más reciente es la combinación sinérgica de las

informaciones procedentes de varios sensores. Así, el espectro visible proporciona mapas de luminancias y reflectancias, y los rangos térmicos, mapas de temperaturas aparentes. En esta línea, se tiende a utilizar de forma simultánea los sistemas visibles y los infrarrojos, cuya combinación proporciona un sistema de visión muy completo, capaz de funcionar tanto de día como de noche.

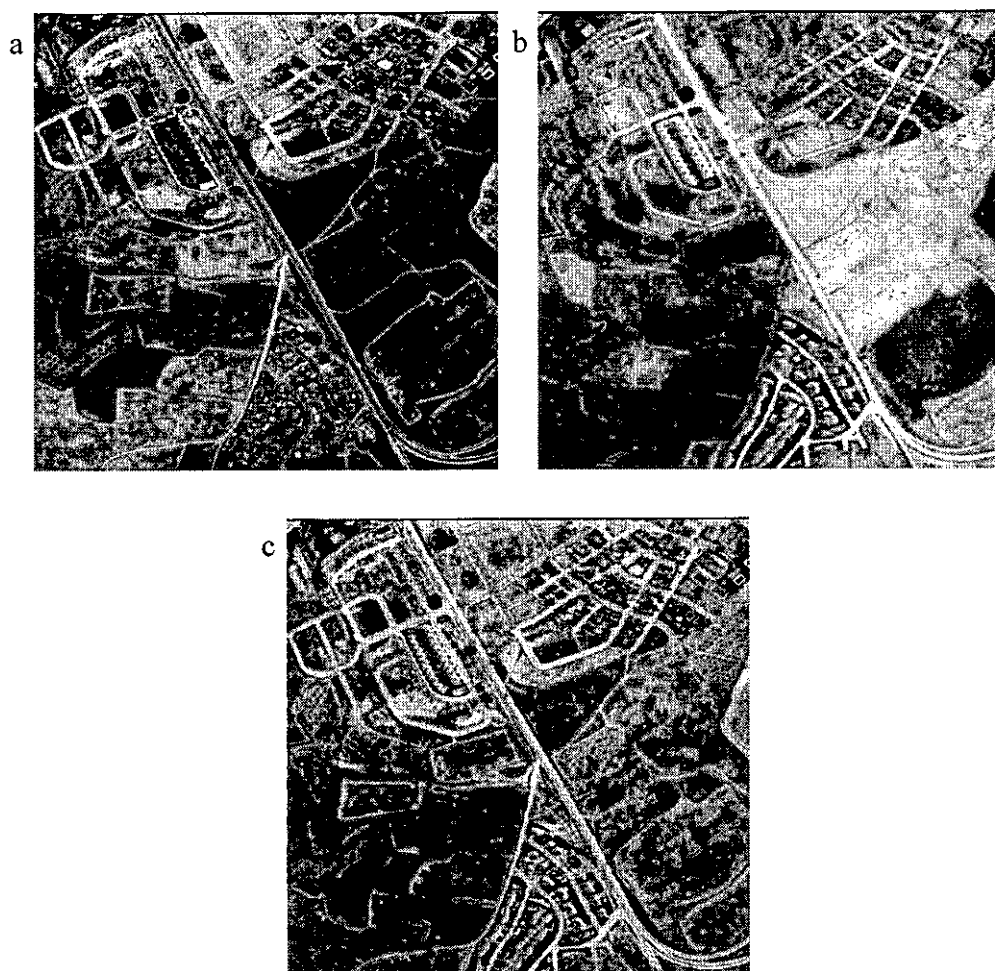
Aunque los sistemas IR son menos sensibles que los sistemas visibles a las condiciones meteorológicas, sus prestaciones también se degradan severamente en situaciones atmosféricas adversas, como elevada humedad, lluvia, o presencia de nubes en la observación aérea. Por tanto, cuando sea necesario disponer de sistemas de observación que no dependan de las condiciones atmosféricas, debe recurrirse a otro tipo de sensores, como, por ejemplo, el Radar de Apertura Sintética (SAR) [2], para el que la atmósfera se presenta prácticamente transparente. La utilización del SAR como sistema aéreo de observación permite, por tanto, obtener información de la superficie terrestre independientemente de cuáles sean las condiciones meteorológicas en la zona a observar. Como contrapartida, las imágenes SAR presentan una resolución menor y unos niveles de ruido mayores que las imágenes visibles o infrarrojas. Por ello, es muy interesante combinar la información procedente de un sistema SAR con la de otros sensores visibles o infrarrojos.

En otros ámbitos se utilizan diferentes métodos para la formación de imágenes. En el campo de la medicina, por ejemplo, se utilizan sistemas que funcionan en otras bandas espectrales, como radiografía (rayos X), termografía (IR), gammagrafía (rayos  $\gamma$ ), tomografía axial computerizada (rayos X), así como otros métodos no basados en radiación electromagnética, como la ecografía, que utiliza ondas sonoras, la tomografía por emisión de positrones (PET), o la resonancia magnética nuclear (RMN), basada en medidas de *spin* nuclear.

Es decir, prácticamente en todos los entornos científicos y tecnológicos se tiende a la utilización de sensores de diferente naturaleza con el fin de obtener información suplementaria que, combinada adecuadamente, permita obtener el máximo conocimiento acerca de los procesos o fenómenos que se desean analizar.

Uno de los métodos existentes y más directos para combinar la información contenida en varias imágenes es la fusión a nivel de píxel. Si bien el método más sencillo para realizar la fusión de varias imágenes consiste en la superposición radiométrica o colorimétrica (falso color) a nivel global, estos métodos no optimizan el contenido local de información, por lo

que su uso es limitado. Recientemente se han desarrollado otros métodos basados generalmente en descomposición por ondículas (*wavelets*), que permiten el análisis multirresolución y multidireccional de la información contenida en las imágenes [3-6]. A modo de ejemplo, en la figura 1.1 se presenta el resultado de la aplicación de un método de fusión [3], basado en la descomposición en ondículas de Gabor [7], a dos imágenes aéreas de Navalcarnero adquiridas mediante el sensor Daedalus DS-1268, en las bandas espectrales de 0.45-0.52  $\mu\text{m}$  y 2.08-2.35  $\mu\text{m}$ . La imagen fundida resultante, 1.1c, contiene todos los detalles presentes en las dos imágenes de entrada, 1.1a y 1.1b.



**Figura 1.1.** Dos imágenes adquiridas mediante el sensor Daedalus DS-1268 en las bandas espectrales (a) 0.45-0.52  $\mu\text{m}$  y (b) 2.08-2.35  $\mu\text{m}$ . (c) Resultado de su fusión, en el que se aprecian detalles presentes en las dos imágenes de entrada. (Imágenes cedidas por el Instituto Nacional de Técnica Aeroespacial, INTA).

En cualquier caso, para poder utilizar de manera conjunta el contenido de las imágenes procedentes de sensores diferentes, bien mediante su fusión a nivel de píxel, a nivel de atributos (*features*), o por cualquier otro procedimiento, es necesario que presenten la



misma distribución espacial, de modo que pueda establecerse una correspondencia única entre los píxeles de unas imágenes y los de otras. Sin embargo, los diferentes sistemas de adquisición, por su distinta naturaleza y configuración, causarán, en el caso más general, diferencias en esa distribución espacial, por lo que será necesario corregirla, deformando las imágenes para su ajuste.

Se denomina *Registro (o Corregistro) de Imágenes* al proceso utilizado para corregir las diferencias en la distribución espacial de dos imágenes de una misma escena, de forma que puntos de iguales coordenadas en las imágenes se correspondan con el mismo punto de la escena que aparece en ellas. Dos imágenes que cumplan este requisito se dice que están *registradas*.

### **1.1. Justificación y propósito**

En el caso de imágenes de la misma banda espectral o de bandas espectrales próximas, el registro puede realizarse mediante técnicas de correlación de imágenes. Sin embargo, la utilización conjunta de sistemas de formación de imagen de naturaleza diferente plantea un nuevo problema de registro, debido a que la información contenida en cada imagen no es correlacionable con la de las otras, puesto que las magnitudes físicas que generan la información son generalmente distintas. Así, por ejemplo, en el registro visible-IR, no se puede poner en correspondencia la reflectancia o la luminancia de un cuerpo con su temperatura aparente. Por ello es necesario disponer de procedimientos que permitan registrar imágenes de cualquier naturaleza.

Los métodos existentes para el registro de imágenes de rangos espectrales próximos no son aplicables en el caso general. Por otro lado, los muy escasos procedimientos descritos recientemente en la bibliografía para el registro de imágenes de bandas espectrales diferentes sólo son aplicables si la corrección de registro es pequeña. Por todo ello, es necesario el desarrollo de nuevos métodos de registro, capaces de corregir imágenes de diferente naturaleza independientemente del grado de desregistro que presenten.

En general, siempre es posible realizar un registro manual de las imágenes, independientemente de cuál sea su origen. Sin embargo, en la mayor parte de los casos, el volumen y las características de las imágenes hacen inviable o de poco interés práctico la utilización de métodos manuales. Por otro lado, en aquellos casos en los que existan

modificaciones frecuentes en la geometría de la observación, posiciones relativas de los sensores, etc, tampoco puede aplicarse el registro manual.

El objetivo fundamental de esta tesis es el desarrollo de procedimientos automáticos de registro aplicables a imágenes de diferente naturaleza. Si bien los procedimientos presentados en esta memoria están concebidos para uso general, se ha puesto especial interés en su aplicación al registro de imágenes visibles, infrarrojas y de radar de apertura sintética (SAR), por ser las de uso más generalizado en nuestra área de trabajo. Ello nos ha permitido tener acceso a imágenes de una misma escena captadas por estos diferentes tipos de sensores.

Los métodos desarrollados para el registro de imágenes de bandas espectrales diferentes son conceptualmente más robustos que los utilizados para el registro de imágenes similares, en cuanto a que deben utilizar información más depurada. Por tanto, la resolución del problema aquí planteado implica a su vez la resolución del problema de registro automático para imágenes similares, que aun siendo más sencillo y existiendo algunos antecedentes, no está resuelto de forma generalizada.

Los procedimientos de registro descritos en esta memoria permiten ampliar de forma considerable el rango de aplicación de campos como, entre otros, reconocimiento automático de objetos en entornos naturales, teledetección (análisis multisensor y multitemporal), mejora y restauración de secuencias de imágenes (mediante la aplicación secuencial de registro, fusión, integración y restauración), configuración de mosaicos de imágenes, etc.

## **1.2. Estructura de la Memoria**

La memoria está organizada de la siguiente forma:

- En el capítulo 2 se describen los conceptos básicos relacionados con el registro de imágenes y se repasan los principales métodos y tipos de registro existentes.
- En el capítulo 3 se estudia qué tipo de atributos o elementos pueden ser empleados para el registro de imágenes de bandas espectrales diferentes y se describen las técnicas necesarias para su detección en las imágenes. Después se explica el procedimiento de

extracción de posibles transformaciones de registro, a partir del casamiento de los conjuntos de atributos detectados.

- El primer procedimiento desarrollado para el registro automático, basado en técnicas de acumulación, se describe en el capítulo 4. Se incluyen ejemplos de su aplicación y se discuten sus limitaciones.
- El segundo, y más depurado, método de registro automático desarrollado, basado en la calidad de casamiento de conjuntos de segmentos lineales rectos, se presenta en el capítulo 5. Se muestran algunos resultados de su aplicación al registro de imágenes de bandas espectrales diferentes.
- La técnica de refinado de la transformación afín de registro para generar transformaciones de órdenes superiores se describe en el capítulo 6. Se incluyen refinamientos de las transformaciones obtenidas en los ejemplos de los capítulos 4 y 5.
- En el capítulo 7 se describen dos formas de evaluar los métodos de registro automático desarrollados, basadas en el análisis de la transformada de Gabor y en la evaluación del campo vectorial de desplazamientos locales de desregistro.
- En el capítulo 8 se presentan algunas de las aplicaciones de las técnicas y métodos de registro desarrollados, en particular, registro visible-IR, registro visible/IR-SAR, detección de formas, generación de mosaicos y registro de mapas e imágenes.
- La valoración final del trabajo realizado, junto con sus posibles aplicaciones y desarrollos futuros, se recoge en el capítulo 9.
- En los apéndices se explica el modelo de transformación afín elegido, y los procedimientos desarrollados para la reducción del ruido de barrido y el crecimiento hexadecagonal de regiones.
- Al final de la memoria se incluye una lista de la bibliografía utilizada.

## 2. Registro de Imágenes

Aunque existen multitud de métodos para registrar imágenes [8], bien generales o sólo para tipos de imágenes específicos, es posible agruparlos en dos o tres clases, dependiendo del grado de intervención humana que precisen para su funcionamiento. Las dos clases principales serían las de registro *manual* y *automático*. También puede hablarse, en algunos casos, de registro *semiautomático*. Esta clasificación se corresponde, además, con el orden histórico en el desarrollo de los diferentes procedimientos de registro existentes.

En la mayoría de las aplicaciones actuales el registro se realiza de forma manual; sólo en algunos casos especiales se dispone de métodos automáticos. En el proceso de registro manual es un operario humano el que identifica parejas de puntos coincidentes en las imágenes a registrar, a partir de las cuales se determinarán las funciones de deformación de una de las imágenes para que quede registrada con la otra.

En el registro automático todo el proceso de registro se realiza sin intervención humana. Hasta ahora sólo puede hablarse de registro automático en algunas aplicaciones concretas, en las que existe un modelo o entorno de laboratorio controlado. Es el caso, por ejemplo, de la estereoscopia, en la que el contenido de las imágenes a registrar es muy similar.

Cuando las imágenes a registrar son muy diferentes, el registro automático resulta complicado, y suele ser necesario recurrir al registro manual. También puede ocurrir que, por su complejidad, sea difícil realizar el registro de forma completamente autónoma, y casi siempre existe algún tipo de umbral que debe ser fijado de antemano por un operador humano. En aquellos casos en los que la supervisión humana sea considerable, puede hablarse de registro semiautomático.

### 2.1. Procedimientos generales de registro

#### 2.1.1. Registro manual

Tradicionalmente el registro de dos imágenes se realiza en tres fases: búsqueda de puntos de control, cálculo de las funciones de deformación y deformación de una de las imágenes para registrarla con la otra. Se suele denominar *imagen fuente* a aquella que será

deformada para que se registre con la otra, a la que se llama *imagen destino*. A continuación se describe en detalle cada una de estas fases:

- Búsqueda de puntos de control

En primer lugar se buscan parejas de puntos correspondientes en las dos imágenes, llamados puntos de control o de anclaje. Los dos puntos de una pareja deben corresponder a un mismo punto de la escena representada en las imágenes. Si las imágenes estuvieran registradas, las coordenadas de cada punto de control en su imagen correspondiente serían iguales, pero, en el caso más general, será necesario deformar una de las imágenes para que los puntos de control coincidan.

Los puntos de control pueden situarse en cualquier lugar de las imágenes, pero lo normal es que se posicionen en lugares característicos, fácilmente identificables en ambas imágenes (normalmente en esquinas de objetos o cruces de líneas). Es muy conveniente que los puntos de control se encuentren homogéneamente distribuidos por toda la imagen, para que las funciones de deformación que se calculen a partir de ellos no se comporten de forma extraña en aquellas zonas donde no se hayan definido puntos de control.

- Cálculo de las funciones de deformación

La deformación de la imagen fuente se determina mediante el análisis de la lista de puntos de control, calculando, mediante un método de minimización, una función de distorsión que ajuste lo mejor posible los puntos de control de la imagen fuente a los de la imagen destino. Si la elección de los puntos de control es correcta y el modelo elegido (transformación afín, polinomios de orden superior, etc) se aproxima bien a la función de deformación real, la distorsión de la imagen fuente establecerá una correspondencia única entre todos los puntos de las imágenes, de modo que puntos de iguales coordenadas en las imágenes se correspondan con el mismo punto de la escena representada en ellas.

Las funciones utilizadas para modelar la deformación suelen ser transformaciones afines, aunque con frecuencia se utilizan polinomios de segundo o tercer orden para compensar las posibles distorsiones de naturaleza no lineal presentes en las imágenes. El ajuste de los coeficientes de las funciones de deformación se realiza por mínimos cuadrados en función de las posiciones de los puntos de control en las imágenes fuente y destino.

- Deformación de la imagen fuente

Utilizando las funciones de distorsión calculadas se deforma la imagen fuente para que quede registrada con la imagen destino. A este procedimiento se le suele denominar *warping*. Al igual que en las transformaciones geométricas normales, el proceso a seguir consiste en la obtención de la función inversa para realizar el remuestreo desde el espacio destino (el de la imagen fuente registrada y el de la imagen destino) hasta el espacio origen (el de la imagen fuente original).

### 2.1.2. Registro automático

En el caso del registro automático la utilización de un procedimiento basado en el emparejamiento de puntos de control presenta varias dificultades. La primera de ellas es que, en el registro manual, los puntos de control se eligen inicialmente con el conocimiento a priori de que existe en la otra imagen un punto correspondiente. El operador humano interpreta las imágenes y es capaz de ir escogiendo en ellas parejas de puntos fuente y destino correspondientes. En el caso del registro automático, la determinación de los puntos de control no es tan sencilla, ya que no tiene por qué suponerse que los puntos de control detectados en una de las imágenes tengan correspondencia en la otra imagen, dado que no existe un proceso de interpretación que los evalúe.

Además, las imágenes de diferente naturaleza presentan grandes diferencias en los niveles de gris, texturas o niveles de ruido, por lo que la utilización de detectores de puntos característicos resulta poco conveniente, dada la variabilidad de su comportamiento frente a estos factores. Otro tipo de atributos (*features*), como bordes o segmentos, pueden resultar más útiles para el registro automático.

Por todo ello, en general, resulta más conveniente descomponer el proceso de registro en las siguientes fases:

- Búsqueda de atributos (*features*)

Inicialmente se buscan, de forma independiente en cada imagen, conjuntos de atributos o elementos que puedan ser útiles para el registro de las imágenes. Estos elementos pueden ser de naturaleza muy variada: esquinas formadas por los bordes extraídos de las imágenes de entrada, parámetros estadísticos basados en regiones de un tamaño determinado, líneas

rectas, etc. Lo importante es que el tipo de atributo elegido pueda ser detectado de forma fiable en las dos imágenes y que esté presente en suficiente número en ambas como para poder ajustar el modelo de deformación utilizado.

- Casamiento (*matching*) de los conjuntos de atributos detectados

Una vez que se han obtenido los conjuntos de atributos en ambas imágenes, es necesario ponerlos en correspondencia, estableciendo relaciones individuales entre elementos fuente y destino que sean consistentes a escala global. Es posible que algunos de los elementos presentes en una de las imágenes no tenga correspondencia en la otra imagen, por lo que el método de casamiento de los conjuntos debe contemplar esta posibilidad.

- Cálculo de las funciones de deformación

En función de los casamientos realizados se aplica el modelo de transformación elegido y se determina la solución óptima para las funciones de deformación de la imagen fuente.

- Deformación de la imagen fuente

La función de deformación calculada se aplica a la imagen fuente para que quede registrada con la imagen destino.

Como puede verse, el procedimiento de registro manual mediante casamiento de puntos de control es un caso particular de esta descripción, más genérica, del proceso de registro de imágenes.

## **2.2. Antecedentes**

En la mayoría de las aplicaciones, el proceso de registro de imágenes obtenidas por medios de adquisición diferentes se fundamenta en un conocimiento previo del tipo de información que contienen. En el caso de las imágenes médicas (por ejemplo, fusión de radiografías con imágenes de resonancia magnética nuclear (RMN)) el registro se apoya en la identificación de las llamadas marcas fiduciales, estructuras que son reconocibles sólo por el especialista. En la geo-referenciación, aunque las imágenes a registrar pueden ser de muy diferente naturaleza, también se parte del conocimiento de la posición de ciertos elementos geográficos. En otros casos, como el reconocimiento de patrones, se dispone previamente

de la descripción de un modelo que es necesario detectar en las imágenes. El registro de ese modelo con las figuras de las imágenes se fundamenta básicamente en técnicas de correlación con máscaras o en la transformada de Hough [9-13]. En algunos casos es posible utilizar marcas externas a los objetos presentes en las imágenes para facilitar su posterior registro. Es el caso del posicionamiento de cruces en las imágenes, aunque también se utilizan marcadores más complejos, como piezas tubulares con compuestos químicos específicos en su interior, en imágenes de RMN [8], por ejemplo.

En el caso más general del registro de imágenes de bandas espectrales alejadas, lo único que puede presuponerse es que ambas corresponden a la misma escena, y que no se va a disponer de ningún tipo de información adicional aparte del propio contenido de las imágenes. En consecuencia, antes de iniciar el proceso de registro, es necesario extraer de las imágenes a registrar toda la información común que sea posible. Según la cantidad de información común que inicialmente esté disponible como base para el registro, puede dividirse el conjunto de los algoritmos de registro en dos tipos principales: algoritmos para el registro de imágenes “similares”, donde las diferencias entre las imágenes son casi exclusivamente de tipo geométrico; y algoritmos para el registro de imágenes “diferentes”, esto es, de bandas espectrales muy alejadas o procedentes de distintos tipos de sensores. En este último caso, además de las diferencias geométricas que es necesario corregir, aparecen otras como variación en los niveles de gris, texturas e incluso diferencias de contenido, causadas por las diferentes respuestas de los objetos presentes en la imagen frente a la radiación o al método utilizado para la formación de las imágenes.

Conviene mencionar que, en principio, cualquier método para el registro de imágenes diferentes debería poder utilizarse para el registro de imágenes similares.

### 2.2.1. Registro de imágenes similares o de bandas espectrales próximas

Si las imágenes a registrar son similares (por haber sido obtenidas con el mismo tipo de dispositivo o ser de bandas espectrales muy próximas) y las deformaciones no son muy grandes, se pueden usar técnicas de correlación [14-20] o de flujo óptico [21-24] para detectar parejas de puntos correspondientes (mediante el análisis de pequeñas ventanas alrededor de los puntos de interés) o para la validación de una determinada transformación (evaluando la superposición de dos imágenes o regiones de imágenes). En estos casos también se pueden utilizar métodos de correlación basados en la transformada de Fourier



[25], que son especialmente útiles porque permiten restaurar de manera sencilla transformaciones rígidas en las imágenes a registrar.

Algunos de los métodos más comunes para estos tipos de registro están recopilados en la referencia [8], aunque existen técnicas más modernas, principalmente para el registro de pares estereoscópicos [26] y de imágenes del mismo sensor [27] o de bandas espectrales cercanas con deformaciones pequeñas [28].

### 2.2.2. Registro de imágenes de bandas espectrales alejadas

Si las imágenes a registrar provienen de diferentes sistemas de formación de imágenes, o del mismo sistema pero funcionando en otra banda espectral, el nivel de respuesta radiométrica (en definitiva, el nivel de gris) y las características de textura de las imágenes no se conservan, y no pueden utilizarse para el registro. Por ello, técnicas genéricas, como la correlación, no resultan útiles en estos casos. Así mismo, es posible que una deformación muy acusada de una de las imágenes no permita la utilización de las técnicas de registro para imágenes similares, aunque las imágenes a registrar sean del mismo tipo.

También pueden aparecer diferencias en los niveles de gris o las texturas si las imágenes han sido capturadas en condiciones de iluminación distintas o instantes temporales diferentes. Es el caso, por ejemplo, de imágenes de la superficie terrestre en las que los cultivos de una zona han sido substituidos por otros. La variación de las condiciones climáticas en una región también puede influir en las imágenes obtenidas por un satélite o avión. Así, la presencia de nubes o un elevado grado de humedad degradan la calidad de las imágenes visibles e infrarrojas, y las variaciones en el nivel de humedad del suelo causan grandes variaciones en la intensidad de la radiación recibida por un sistema de formación de imagen basado en SAR.

No obstante, existen algunas características que tienden a conservarse en imágenes de diferente naturaleza, como las esquinas o los bordes de los objetos presentes en las imágenes. Aunque un salto térmico (en una imagen IR) o un salto de color (en una imagen visible) no tiene por qué corresponderse con el borde físico de un objeto de la escena, un borde físico presente en la escena tiene muchas probabilidades de causar un salto en luminancia, temperatura aparente o cambio de estructura superficial. Por ello puede decirse que los bordes físicos tienden a percibirse como tales en las imágenes, independientemente

del método de adquisición empleado, y pueden ser utilizados como elementos invariantes para el registro de imágenes de bandas espectrales alejadas.

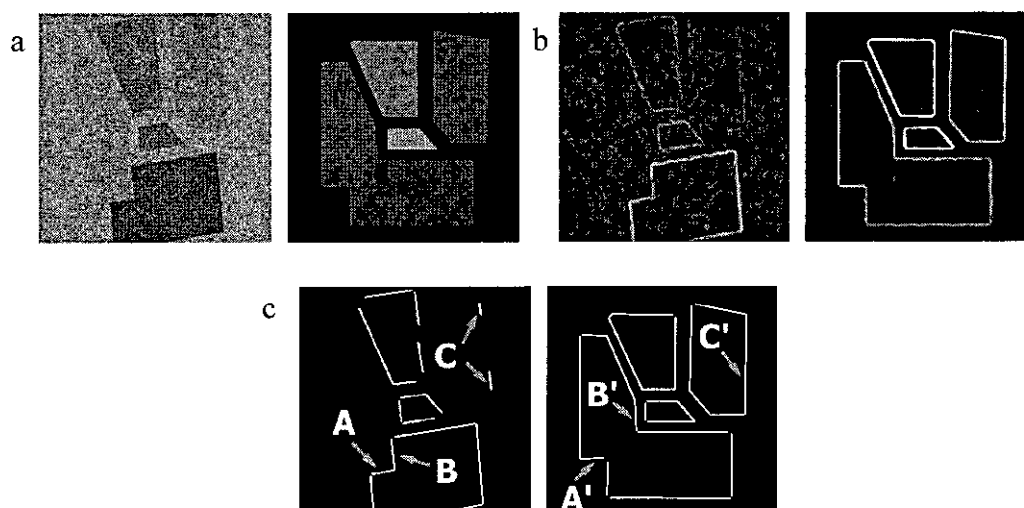
Por otro lado, la información contenida en las imágenes de bordes (obtenidas mediante la aplicación de un filtro de extracción de bordes) es menor y más dispersa, en general, que la existente en una imagen de grises, con lo que las técnicas de registro empleadas para el registro de imágenes similares no pueden aplicarse. Así, aunque existen algunos métodos de casamiento (*matching*) basados en correlación de bordes para la detección de objetos o formas [29-30], la aplicación de este tipo de técnicas al registro de imágenes está limitada, casi exclusivamente, a la minimización del error cuando el sistema ya se encuentra muy cerca de la transformación de registro óptima, debido a que la respuesta de la correlación de bordes, lejos de ese punto, varía de forma prácticamente aleatoria.

Las esquinas (esto es, puntos de los bordes que tengan una curvatura elevada) y otros tipos de puntos característicos (*feature points*) obtenidos mediante la aplicación de algún operador [28-31] pueden usarse para visión estereoscópica o para el registro de imágenes similares. Su aplicación para el registro de imágenes de bandas espectrales alejadas, sin embargo, presenta varios problemas. En primer lugar, resulta difícil diseñar un método de extracción de esquinas que se comporte de la misma manera en imágenes de distinta naturaleza. Por otra parte, los métodos de registro basados en puntos característicos son muy sensibles a las diferencias de escala, lo cual, unido a la dificultad de asegurar las correspondencias entre puntos, hace a estos métodos poco robustos, excepto para la reconstrucción de deformaciones pequeñas. En la referencia [32], prácticamente la única que aplica este tipo de técnicas (en concreto la descrita en [28]) al registro de imágenes de bandas espectrales separadas (visible-IR), ya se ponen de manifiesto los problemas que existen para registrar imágenes con diferencias de escala mayores del 10%.

Otros atributos que pueden utilizarse para el registro son segmentos rectos o *splines* ajustados a los bordes de las imágenes. El registro basado en el casamiento de segmentos rectos se aplica de forma extensiva en visión estereoscópica [29, 33-41], donde se puede suponer que ciertas características de los segmentos se conservan entre las imágenes a registrar. Principalmente se suele utilizar como medida de valoración de casamientos la superposición entre segmentos, que implica la conservación de las coordenadas relativas de los extremos de los segmentos o su longitud. Estos atributos, sin embargo, no se preservan entre imágenes de diferentes bandas espectrales, y debe explotarse alguna otra característica que se conserve en imágenes de diferente naturaleza.

El método de registro propuesto en esta tesis únicamente supone que las rectas definidas por los segmentos se conservan entre las imágenes a registrar, una condición mucho más débil, pero más fácil de garantizar, que las anteriores. Esta condición se fundamenta en el hecho de que un borde físico presente en la escena puede dar lugar a diferentes configuraciones de segmentos en las imágenes captadas por diferentes sensores.

Un ejemplo de este fenómeno puede verse en la figura 2.1, en la que se muestra un par IR-visible de imágenes sintéticas, generado a partir de una misma imagen base (la “escena”) alterando sus niveles de gris y añadiéndole distintos niveles de ruido, para simular las diferencias que se observan entre pares IR-visibles reales. La imagen IR sintética presenta, además, una deformación afín respecto de la imagen visible, para simular un desregistro. La aplicación de un extractor de bordes a las imágenes genera resultados distintos (figura 2.1b), lo cual se traduce finalmente en conjuntos de segmentos detectados diferentes (2.1c). Así, por ejemplo, los segmentos marcados A y A' se corresponden con el mismo borde físico de la escena y, por tanto, casan, aunque no se superpongan cuando se aplique la transformación de registro. Lo mismo les sucede a los segmentos B y B'. Los dos segmentos marcados C casan con el mismo segmento C', debido a que el procedimiento de extracción de segmentos no ha sido capaz de detectar el segmento C completo.



**Figura 2.1.** Comportamiento de los detectores de bordes y segmentos sobre imágenes de una misma escena en diferentes bandas espectrales. (a) Simulación de un par de imágenes IR-visible. (b) Bordes extraídos y (c) segmentos detectados. A-A', B-B' y C-C' indican segmentos correspondientes a los mismos bordes físicos de la escena pero que presentan configuraciones diferentes después del proceso de extracción.

La aparición de estas diferencias entre conjuntos de segmentos extraídos de imágenes de bandas espectrales alejadas, invalida todos los métodos basados en conservación de

longitudes, conservación de las coordenadas de los vértices, superposición y demás técnicas empleadas para el registro de imágenes similares o pares estereoscópicos.

Sin embargo, la condición de casamiento basada en la pertenencia a la misma recta (correspondiente, en última instancia, al borde físico real existente en la escena), es perfectamente aplicable y puede emplearse como fundamento para el desarrollo de métodos de registro de imágenes de diferentes bandas espectrales. Este es el principio básico sobre el que hemos desarrollado los procedimientos de registro descritos en esta memoria.



### 3. Extracción de estructuras geométricas y determinación de transformaciones

En este capítulo se estudian los atributos o estructuras que pueden emplearse para el registro de imágenes de bandas espectrales diferentes y los procedimientos necesarios para su detección en las imágenes a registrar. Finalmente, se describe el método de cálculo de posibles transformaciones de registro globales a partir de casamientos locales entre los atributos o elementos elegidos.

#### 3.1. Elementos invariantes

Prácticamente los únicos atributos que pueden preservarse entre imágenes de diferente naturaleza son bordes de los objetos presentes en la escena. Es posible, entonces, que algunos subconjuntos de elementos o atributos contenidos en los bordes, como segmentos rectos o esquinas, se conserven también y puedan utilizarse para el registro de imágenes diferentes. Los bordes extraídos de las dos imágenes a registrar, no obstante, no suelen ser exactamente iguales. Las diferencias pueden estar causadas por el distinto comportamiento de los extractores de bordes, que se ven afectados por los diferentes niveles de ruido o contraste de las imágenes, aunque también puede ocurrir que los contornos de los objetos presentes en las imágenes no sean, efectivamente, iguales, debido a que las magnitudes físicas representadas en las imágenes son diferentes. Es necesario, por tanto, tener en cuenta esas diferencias para el diseño de un método de registro robusto.

La utilización de esquinas o, más generalmente, puntos característicos (*feature points*), para el registro ya ha sido propuesta con anterioridad [28, 31, 32, 42, 43]. Sin embargo, los detectores de puntos característicos suelen ser muy sensibles a las diferencias de escala o sesgo (*skew*) y algunos de ellos (principalmente los que trabajan directamente sobre la imagen de niveles de gris) también a la rotación, el nivel de ruido o las diferencias de textura, y pueden generar conjuntos de puntos muy diferentes en dos imágenes de una misma escena adquiridas en diferentes bandas espectrales. Los extractores de esquinas que funcionan sobre representaciones vectorizadas de los bordes (cadenas de píxeles, segmentos) se ven afectados por las diferencias existentes en los bordes extraídos (fragmentación y falta de coincidencia espacial, principalmente), y también generan diferencias en los conjuntos de puntos resultantes.

Esas diferencias entre los conjuntos de puntos a registrar dificultan el subsiguiente proceso de establecimiento de correspondencias entre elementos de los dos conjuntos. Las técnicas de casamiento de conjuntos de puntos, en general, tampoco son muy robustas frente a deformaciones geométricas severas, por lo que sólo pueden utilizarse cuando la deformación a corregir es pequeña o se dispone de algún tipo de prerregistro previo, como se discute en el artículo de Mason y Wong [41].

Una opción más robusta consiste en utilizar correspondencias entre segmentos rectilíneos contenidos en los bordes, ya que estos son más estables que los puntos característicos frente a las variaciones presentes en imágenes diferentes. Algunos autores ya los han utilizado para el registro de pares estereoscópicos o reconstrucción 3D [33-37, 44], donde la semejanza de las imágenes a registrar se refleja en los conjuntos de segmentos extraídos, que suelen ser muy similares. La aplicación de esas técnicas para el registro de imágenes diferentes presenta, sin embargo, algunos problemas, debidos a las diferencias existentes entre dichos conjuntos.



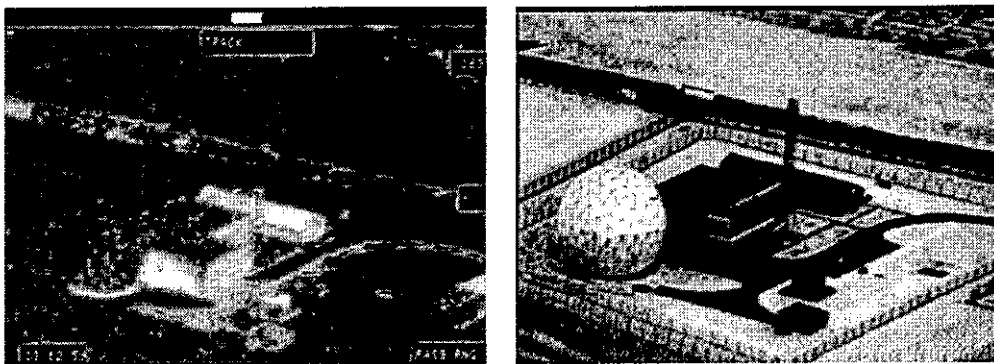
**Figura 3.1.** Diferencias en los conjuntos de segmentos extraídos de un par de imágenes de una misma escena tomadas en diferentes bandas espectrales. Fila superior: imagen IR, bordes extraídos y segmentos detectados; fila inferior: imagen visible, bordes y segmentos.

El primer problema es que, con frecuencia, los segmentos aparecen fragmentados, incompletos o pueden no aparecer en la misma posición en una de las imágenes. Este fenómeno se muestra en la figura 3.1, donde puede verse que prácticamente todos los bordes detectados en las imágenes presentan algún grado de fragmentación o

incompletitud, y cómo segmentos correspondientes de las dos imágenes rara vez tienen las mismas posiciones o longitudes relativas.

Por lo tanto, atributos como la superposición de los segmentos, su longitud o la coincidencia entre sus vértices (aunque son útiles para el registro de imágenes similares [29, 33-41]) no se pueden garantizar y no deben utilizarse para el registro de imágenes diferentes. Otras características asociadas a los segmentos, como el contraste entre las regiones adyacentes al segmento, también útiles para imágenes similares [33, 42], tampoco pueden utilizarse cuando se trabaja con imágenes procedentes de distintos tipos de sensores. El fenómeno de inversión de contraste aparece con frecuencia entre imágenes de diferentes bandas espectrales, como puede verse en la figura 3.2, en donde algunos objetos (p.ej. las carreteras) que aparecen claros sobre fondo oscuro en una de las imágenes, lo hacen a la inversa en la otra. Es más, en los sistemas IR, la polaridad suele ser configurable por el usuario.

Para el desarrollo de los métodos de registro expuestos en esta memoria, se ha tomado como premisa que la única condición razonable para establecer la correspondencia entre dos segmentos de imágenes diferentes es que provengan del mismo borde físico de la escena. Y, puesto que el borde puede aparecer fragmentado o incompleto de forma diferente en las dos imágenes, sólo puede esperarse que los dos segmentos estén contenidos en una misma recta, correspondiente a la proyección, sobre el plano de imagen, del borde físico de la escena del que proceden.



**Figura 3.2.** Una imagen IR (izquierda) y otra visible (derecha) de la misma escena, en la que puede apreciarse claramente el fenómeno de inversión de contraste. Las carreteras y algunos de los edificios, que aparecen claros sobre un fondo más oscuro en la imagen IR, se ven oscuros sobre fondo claro en la visible.



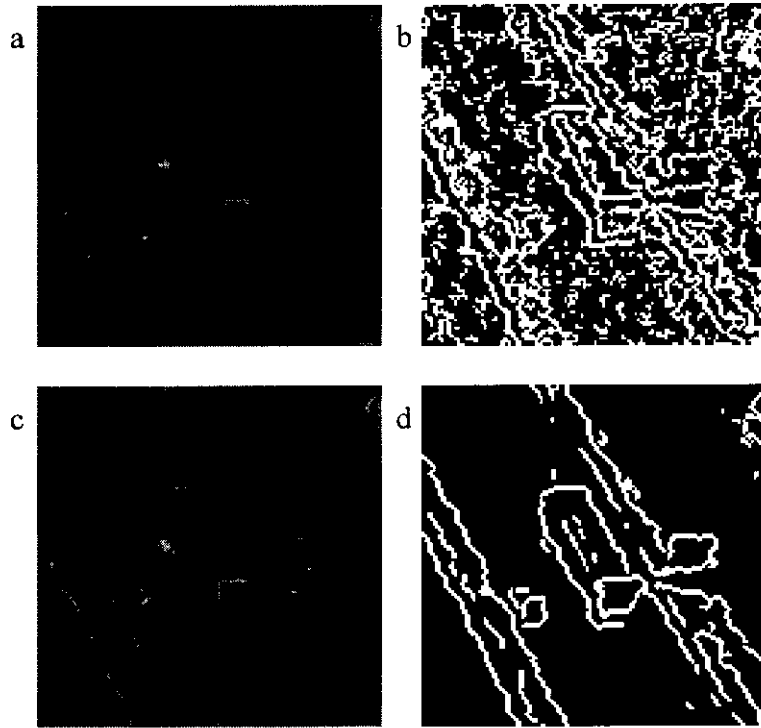
## 3.2. Preprocesado de las imágenes

La extracción de los bordes de las imágenes es una parte clave del procedimiento de registro automático, puesto que determina los atributos que serán utilizados posteriormente para el cálculo y la evaluación de las posibles transformaciones de registro. Por ello, es importante que el método utilizado para la extracción de los bordes genere resultados similares en las dos imágenes a registrar, aunque es de esperar que las diferencias existentes entre ellas repercutan en los bordes obtenidos. Es importante además que, a partir de las imágenes de bordes, puedan extraerse conjuntos de segmentos parecidos, ya que la transformación de registro se determina mediante el establecimiento de correspondencias entre ellos.

Si las imágenes a registrar presentan niveles de ruido o de contraste muy diferentes, será necesario realizar algún tipo de procesado para poder aplicar la misma configuración del extractor de bordes elegido a ambas imágenes. Las imágenes infrarrojas, por ejemplo, suelen presentar un mayor nivel de ruido y menor definición que las imágenes visibles, aunque su calidad ha mejorado mucho últimamente, gracias al desarrollo de nuevos detectores infrarrojos basados en FPAs (*Focal Plane Arrays*).

### 3.2.1. Filtrado de ruido

El ruido presente en las imágenes a registrar puede dificultar el proceso de extracción de los bordes o generar bordes espurios. Para el filtrado de ruido suele ser suficiente con utilizar un filtro gaussiano o uno de mediana [9-13], aunque imágenes muy degradadas pueden requerir procesados más complejos. El principal problema de la convolución con una gaussiana es que también suaviza los bordes útiles de la imagen, y algunos de ellos pueden no ser detectados en el subsiguiente proceso de extracción, sobre todo si el nivel de ruido es alto y la gaussiana debe ser grande. Por ello, en imágenes muy ruidosas, suele ser más conveniente utilizar filtrados que conserven los bordes, como los basados en difusión anisótropa [45] o relajación paramétrica [46]. Hemos utilizado este último método para el filtrado de imágenes de baja calidad. En la figura 3.3 se muestra la diferencia en los resultados de la extracción de bordes con o sin filtrado previo del ruido de la imagen IR.

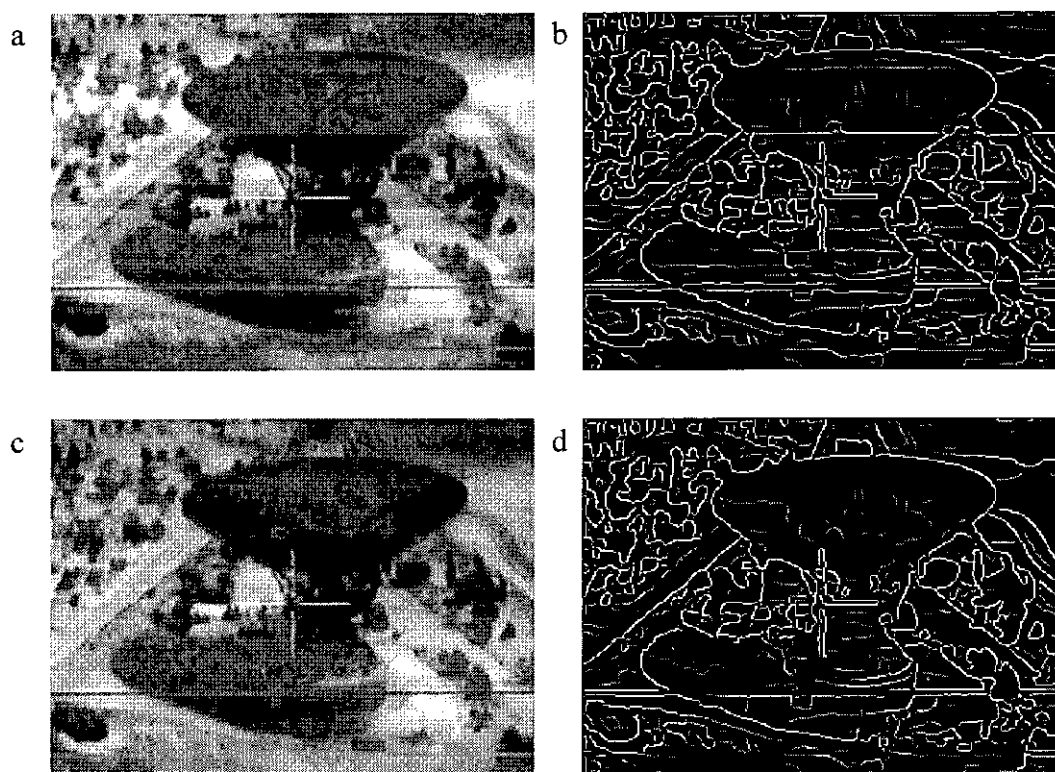


**Figura 3.3.** (a) Imagen IR original, (b) bordes detectados, (c) imagen IR filtrada mediante relajación paramétrica y (d) bordes extraídos de la imagen filtrada con el mismo extractor de bordes utilizado para obtener (b).

El bajo contraste de algunas imágenes también puede afectar al proceso de extracción, de modo que no se obtengan bordes suficientes como para generar un conjunto de segmentos utilizable. En estas situaciones, la aplicación de una rampa [9-12] al histograma de la imagen suele ser suficiente, aunque a veces se obtienen mejores resultados con una ecualización del histograma [9-12].

Existen otros fenómenos que también pueden influir negativamente sobre los bordes extraídos de las imágenes IR. Es el caso de las diferencias locales de contraste que aparecen en imágenes obtenidas mediante dispositivos que funcionan por barrido sobre un vector de detectores IR, y que se manifiestan como bandas horizontales en las imágenes. La presencia de estas bandas introduce bordes espurios y puede causar la fragmentación de algunos bordes que las atraviesen. Muchas de las imágenes disponibles cuando se inició este trabajo procedían de detectores de barrido, por lo que fue necesario desarrollar un método de reducción del bandeo, que se describe en el apéndice B. En la figura 3.4 se muestra el resultado de su aplicación a una imagen IR de barrido, en la que puede observarse cómo el filtrado mejora el funcionamiento del subsiguiente proceso de extracción de bordes.

El barrido de los detectores en las cámaras IR también puede generar deformaciones geométricas en las imágenes. Si se dispone de un modelo adecuado para representarlas, pueden restaurarse directamente mediante el registro con otra imagen, pero lo más general es que las deformaciones sean de naturaleza local y presenten variabilidad temporal (debida, por ejemplo, a movimientos de la plataforma del sensor). En ocasiones, sin embargo, es posible restaurar algunas de estas deformaciones mediante consideraciones relativas al método de adquisición (por ejemplo, distorsiones debidas al barrido en imágenes obtenidas en vuelos aéreos, que pueden compensarse en función de mediciones sobre las perturbaciones de la trayectoria). En esos casos conviene restaurarlas previamente, para simplificar el modelo utilizado en el proceso de registro y agilizar su funcionamiento.



**Figura 3.4.** (a) Imagen IR original y (b) bordes extraídos. (c) Resultado de la aplicación del algoritmo de reducción del bandeo a la imagen original y (d) bordes resultantes.

### 3.2.2. Extracción de bordes y detección de segmentos lineales

El proceso de detección de los segmentos lineales rectos, que se utilizarán como elementos básicos en la determinación de posibles transformaciones de registro, se realiza en tres pasos principales:

- Extracción de bordes en las imágenes a registrar.
- Binarización y adelgazamiento de las imágenes de bordes.
- Vectorización de las cadenas de píxeles de las imágenes de bordes mediante aproximación por listas de segmentos lineales rectos.

A continuación se describen estos pasos en detalle.

#### 3.2.2.1. Extracción de bordes

En general, si las imágenes de entrada no presentan un elevado nivel de ruido o diferencias de contraste grandes, puede utilizarse cualquier extractor de bordes [9-13]. Sin embargo, ésa no es la situación normal, y suele requerirse algún tipo de filtrado previo. Por ello, detectores como el de Canny [47] o el de bordes difusos [48] resultan especialmente interesantes, ya que implícitamente realizan un filtrado antes de la extracción. Pueden, además, funcionar a diferentes escalas y, por tanto, ser utilizados como base para la aplicación de técnicas multirresolución.

El filtro de Canny emplea una derivada de gaussiana como máscara de convolución para la detección de los bordes de una imagen, a la que previamente se aplica un filtrado gaussiano de reducción de ruido. Por su parte, el extractor de bordes difusos utiliza un doble filtrado, de media y de mediana, para generar una imagen de gradientes a partir de la cual se obtienen las cadenas de píxeles correspondientes a los bordes.

En nuestro caso, y dada la importancia del proceso de extracción de bordes para garantizar el buen funcionamiento del método de registro, hemos desarrollado una versión multirresolución del filtro de Canny que permite obtener bordes a cualquier escala pero sin la pérdida de localización asociada a la convolución con la gaussiana inherente al propio filtro.



**Figura 3.5.** Comparación de la implementación usual del filtro de Canny (columna izquierda) y la implementación multirresolución propuesta (columna derecha), para anchuras de gaussiana,  $\sigma$ , (de arriba abajo) de 0.5, 5.0 y 10.0 píxeles.

En nuestra implementación se realiza inicialmente una extracción de bordes a la escala deseada,  $\sigma$ . Los bordes detectados se binarizan y se dilatan. Después se realiza otra extracción de bordes a una escala inferior, y los bordes que estén fuera de las zonas cubiertas por los bordes dilatados en la etapa anterior son descartados. El proceso se repite hasta llegar a una escala mínima fijada de antemano,  $\sigma_{min}$ . Se consigue, de este modo, obtener los bordes relevantes a la mayor escala empleada, pero con el detalle correspondiente al filtro de menor anchura. En la figura 3.5 pueden compararse los resultados obtenidos con el filtro de Canny y con nuestro método multirresolución ( $\sigma_{min} = 0.5$  píxeles), ajustados a diferentes escalas ( $\sigma = 0.5, 5.0$  y  $10.0$  píxeles).

El correcto funcionamiento del proceso de dilatación de las imágenes binarias, que genera la máscara entre cada dos etapas del proceso de extracción de bordes multirresolución, es crucial para la obtención de un buen resultado final. Por ello, y debido a la anisotropía que presentan la mayoría de los principales métodos de dilatación, hemos desarrollado una técnica para el crecimiento hexadecagonal de regiones [49] que genera dilataciones más próximas al crecimiento euclídeo ideal. El método se basa en la aplicación selectiva de dilataciones en 8 y 4 vecinos, y en la inhibición del crecimiento de los píxeles vértice en determinadas etapas del proceso de dilatación. En el apéndice C se presenta una descripción detallada de esta técnica de crecimiento.

### **3.2.2.2. Binarización y adelgazamiento**

Una vez obtenidos los bordes de las imágenes, es necesario binarizarlos y adelgazarlos para obtener cadenas de píxeles de anchura unidad, que serán posteriormente aproximadas por listas de segmentos lineales rectos. El proceso de binarización se realiza por umbralización. El nivel de gris umbral es un parámetro del proceso que debe ser predefinido, aunque es posible utilizar técnicas de umbralización automática [9-13]. La elección de un umbral elevado permite trabajar únicamente con los bordes más destacados de las imágenes y suele ser suficiente para realizar su registro de primer orden. Para el refinado de orden superior de la transformación puede ser necesario recalcular la imagen binaria con un umbral más bajo, de forma que se tengan en cuenta más bordes. La elección de estos umbrales de binarización dependerá del tipo de imágenes a registrar, y debe realizarse antes de iniciar el proceso de registro.

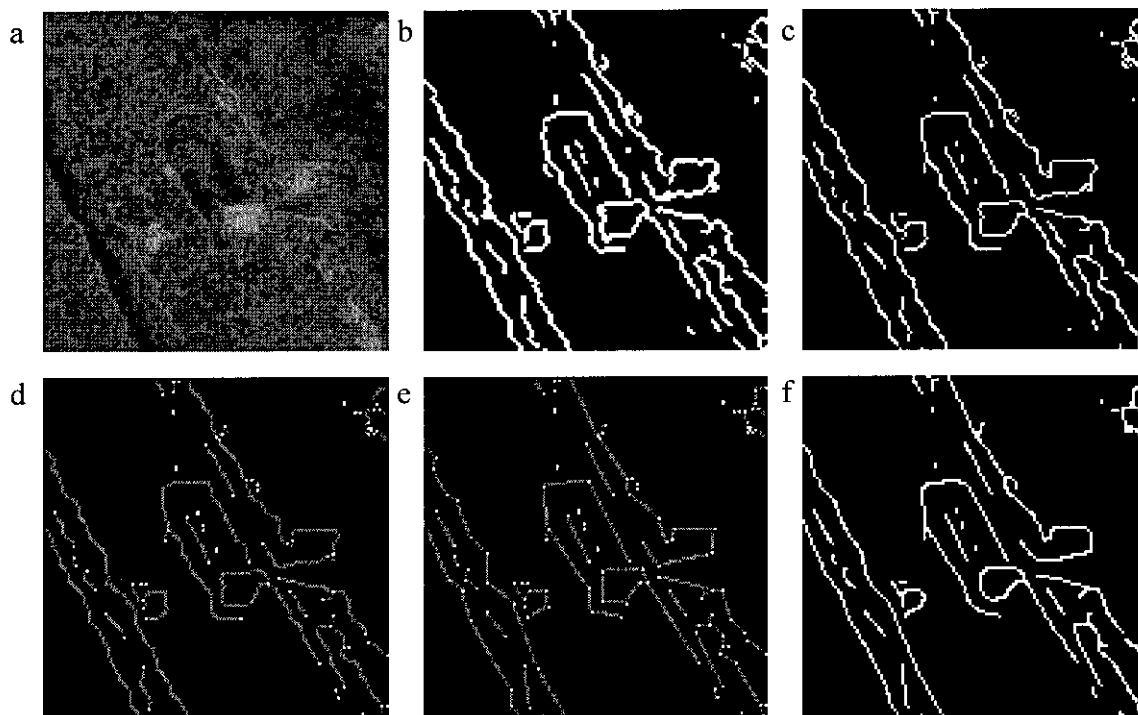
Para el adelgazamiento de las imágenes de bordes binarizados se ha empleado el método descrito en el libro de Pratt [11], que se basa en la erosión selectiva de los píxeles del contorno de las regiones de modo que no se altere su topología (no se fragmenten ni se creen agujeros) ni se erosionen píxeles con un solo vecino. El proceso de adelgazamiento reduce la anchura de los bordes a un píxel, lo que facilita su posterior segmentación mediante el seguimiento de las cadenas de píxeles conexos.

### **3.2.2.3. Vectorización de las cadenas de píxeles**

El proceso de extracción de las cadenas de píxeles comienza con la clasificación de los puntos de la imagen de bordes binarizados y adelgazados, en función del número de vecinos que tengan. Los píxeles aislados se descartan, y los que tienen uno o más de dos vecinos se marcan como extremos de cadenas. Después se siguen las cadenas de píxeles

conexos entre puntos extremos y se almacenan en una lista. Las cadenas se vectorizan entonces mediante el ajuste de segmentos lineales rectos.

Para la vectorización utilizamos el método de Ramer [10], que ajusta recursivamente segmentos a la cadena hasta que la distancia máxima desde los píxeles de la cadena al segmento más cercano se hace inferior a un umbral prefijado. Este umbral debe ser lo suficientemente grande como para evitar que el ruido presente en los bordes de la imagen no fragmente excesivamente las cadenas de píxeles, pero no tanto que se pierdan los detalles de la estructura de los bordes. Para las imágenes de prueba de los algoritmos de registro mostradas en esta memoria se han utilizado normalmente umbrales de 1 a 2 píxeles.



**Figura 3.6.** (a) Imagen test, (b) bordes extraídos, (c) bordes adelgazados, (d) cadenas de píxeles conexos detectadas, (e) aproximación de las cadenas mediante listas de segmentos y (f) ajuste de *B-splines* cúbicos a las listas de segmentos.

Hemos probado también otras representaciones para los bordes de las imágenes, como, por ejemplo, la aproximación de las listas de segmentos encadenados mediante *B-splines* cúbicos [50]. Esta representación puede utilizarse después para la detección de puntos de curvatura elevada, correspondientes a esquinas presentes en las imágenes (y que pueden usarse como puntos característicos), o para suavizar los cambios bruscos de orientación inherentes a las listas de segmentos rectos encadenados. Sin embargo, algunas de las

---

esquinas definidas por segmentos pequeños pueden verse suavizadas excesivamente o desplazadas de su posición original, por lo que finalmente no los hemos utilizado.

En la figura 3.6 se muestran los resultados de los procesos de adelgazamiento, extracción de cadenas de píxeles, aproximación de las cadenas mediante listas de segmentos rectos y ajuste de B-splines cúbicos a los segmentos. En la figura 3.1 también pueden verse los bordes binarizados extraídos de un par de imágenes IR-visible, y los segmentos rectos detectados.

### **3.3. Transformaciones asociadas a casamientos locales de estructuras geométricas**

Para la determinación de posibles transformaciones de registro es necesario establecer relaciones entre los segmentos detectados en las imágenes a registrar. Para ello es conveniente agruparlos en estructuras tales que el casamiento de una estructura de la imagen fuente con otra de la imagen destino defina una transformación candidata para el registro de las imágenes. En este sentido, hemos desarrollado métodos de registro basados en agrupaciones de segmentos en tríos y parejas, así como para la utilización directa de segmentos aislados. Los casamientos entre parejas de segmentos y entre segmentos aislados, sin embargo, no permiten determinar directamente los seis parámetros de una transformación afín completa, por lo que deben utilizarse en combinación con las técnicas de acumulación descritas en el capítulo 4.

#### **3.3.1. Casamiento de triángulos**

Algunos autores ya han desarrollado o propuesto métodos que emplean triángulos o tríos de segmentos para el registro de imágenes [35, 41, 51, 52]. En la referencia [51] se propone la acumulación de las posibles transformaciones de registro asociadas a la correspondencia de diferentes atributos (*features*), pero sólo se restauran la rotación y la traslación. Además, la determinación de las posibles transformaciones en función de las correspondencias se basa últimamente en la superposición de los elementos para su cálculo, con lo que no es posible aprovechar ese método para el registro de imágenes de bandas espectrales alejadas.

En la referencia [52] se usa el casamiento de triángulos para estimar la orientación y la posición de una forma plana en una imagen. De nuevo, solamente se restauran el giro y la



traslación, y además el método únicamente puede utilizarse para la detección de una forma aislada. Extrañamente, la determinación de los posibles triángulos no se realiza en función de los bordes de la figura, sino en base a la unión de puntos obtenidos mediante intersecciones del contorno de la figura con círculos, lo cual complica aún más el método.

Las referencias [35] y [41] presuponen la conservación de las longitudes y posiciones relativas de los segmentos y de sus puntos extremos, condiciones que no se verifican cuando se trabaja con segmentos extraídos de imágenes de diferente naturaleza.

En nuestro método de registro, empleamos el casamiento entre triángulos de la imagen fuente con triángulos de la imagen destino para determinar posibles transformaciones de registro globales. Los conjuntos de triángulos de las imágenes fuente y destino se obtienen agrupando los segmentos detectados en grupos de tres. Cada tres segmentos no colineales, y cuyas prolongaciones no se corten en el mismo punto, definen un triángulo. El triángulo resultante puede determinarse aun cuando los segmentos no estén contenidos en los lados del mismo, como se muestra en la figura 3.8.a.

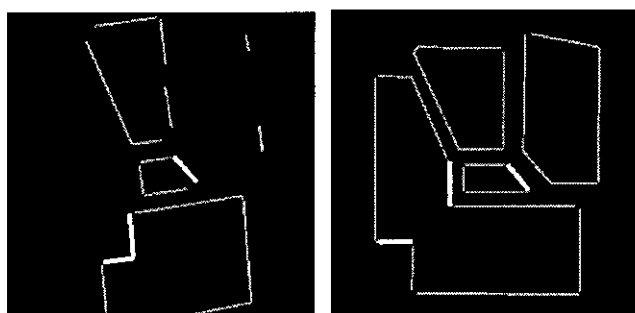
El número de triángulos,  $N_T$ , que se obtienen para un conjunto de  $N$  segmentos detectados en una imagen es del orden de:

$$N_T = C_{N,3} = \frac{N!}{3!(N-3)!} = \frac{N \cdot (N-1) \cdot (N-2)}{6} \approx N^3$$

aunque muchos de los triángulos de la lista pueden ser descartados mediante consideraciones geométricas. Así, durante el proceso de generación de triángulos se descartan aquellos que tengan un área muy pequeña o un ángulo muy agudo, puesto que su forma es muy inestable frente a pequeñas alteraciones en la orientación de los segmentos que los forman. El área y el ángulo mínimos son parámetros configurables del algoritmo de registro, pero no son críticos para su funcionamiento, ya que sólo permiten descartar triángulos que casi con toda seguridad no darán lugar a transformaciones de registro útiles. En las pruebas de registro realizadas durante el desarrollo de esta memoria se han utilizado normalmente 16 píxeles para el área mínima y 0.018 radianes (aproximadamente  $1^\circ$ ) para el ángulo mínimo. También pueden eliminarse redundancias en la lista de triángulos descartando las contribuciones de segmentos colineales, como se describe en la sección 5.2.3.

La principal ventaja de agrupar los segmentos para formar triángulos es que el casamiento de un triángulo de la imagen fuente con uno de la imagen destino define directamente una transformación afín completa (rotación, escalado, sesgo y traslación), que constituye una posible transformación de registro entre la imagen fuente y la imagen destino. Estos triángulos, no obstante, deberían llamarse “triángulos virtuales” ya que, en general, no se corresponderán con estructuras triangulares reales presentes en las imágenes.

Otra ventaja es que dos triángulos siempre pueden hacerse corresponder, incluso cuando los segmentos que los componen no se superpongan una vez que las imágenes hayan sido registradas, como se muestra en la figura 3.7. Obsérvese como la transformación que registra los triángulos sombreados también registra globalmente las imágenes, aunque algunos de los segmentos que definen los triángulos no se superpondrán con su segmento correspondiente en la otra imagen cuando se aplique la transformación de registro. Es el caso de los segmentos A y A' (etiquetados en la figura 2.1c), que no se superponen pero pertenecen al mismo borde físico existente en la escena.

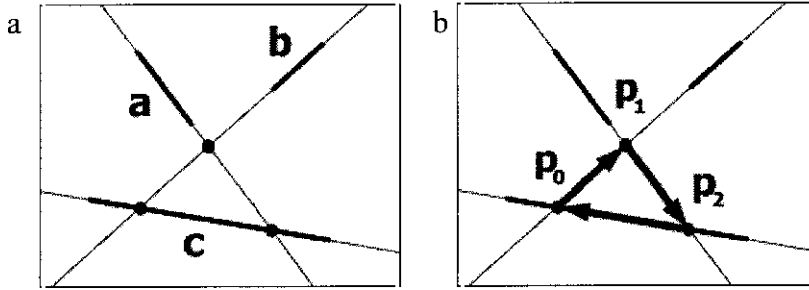


**Figura 3.7.** Triángulos correspondientes (gris oscuro) en las imágenes de segmentos del par de imágenes de la figura 2.1. Los segmentos que definen los triángulos se han resaltado en blanco. La transformación que registra los triángulos también registra las imágenes completas.

Los triángulos son, por tanto, elementos muy útiles en la determinación de posibles transformaciones de registro para imágenes de bandas espectrales alejadas, ya que no requieren la conservación de longitudes, posiciones relativas, vértices o cantidad de superposición de los segmentos que los componen, como ocurre con casi todos los algoritmos de registro para imágenes similares basados en bordes.

Para la determinación del triángulo se calculan los tres puntos de corte entre las rectas definidas por la prolongación de los tres segmentos. Una vez obtenidos, se ordenan según el sentido de las agujas del reloj, empezando por uno cualquiera de ellos (figura 3.8). El orden en el recorrido de los vértices del triángulo debe ser preservado por la

transformación de registro, siempre y cuando no se permitan reflexiones especulares de las imágenes, que suele ser el caso más habitual en el registro de imágenes.



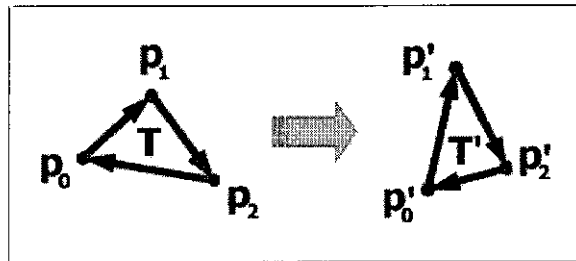
**Figura 3.8.** (a) Un grupo de tres segmentos  $\{a, b, c\}$  definen un triángulo, aunque los segmentos no estén contenidos en los lados del mismo. (b) Una vez que los puntos de intersección de las rectas definidas por los segmentos han sido determinados, se etiquetan con subíndices crecientes según el sentido de las agujas del reloj.

En la figura 3.9 se muestra el casamiento entre un triángulo fuente  $T = \{p_0, p_1, p_2\}$  y uno destino  $T' = \{p'_0, p'_1, p'_2\}$ . Existen tres formas de asignar los vértices de  $T$  a los de  $T'$  preservando el orden de recorrido de los vértices:

$$\left. \begin{aligned} \{p_0, p_1, p_2\} &\rightarrow \{p'_0, p'_1, p'_2\} \\ \{p_0, p_1, p_2\} &\rightarrow \{p'_1, p'_2, p'_0\} \\ \{p_0, p_1, p_2\} &\rightarrow \{p'_2, p'_0, p'_1\} \end{aligned} \right\} \Leftrightarrow \{p_0, p_1, p_2\} \rightarrow \{p'_k, p'_{(k+1) \bmod 3}, p'_{(k+2) \bmod 3}\} \quad (3.1)$$

$$k = 0, 1, 2$$

Si fuese necesario tener en cuenta las reflexiones especulares, bastaría con considerar, además de estas asignaciones, las otras tres asociadas al sentido de giro inverso.



**Figura 3.9.** Correspondencia entre un triángulo fuente y uno destino. Según el orden de asignación de vértices se obtendrá cada una de las tres posibles transformaciones de registro.

Cada una de las asignaciones,  $k = 0, 1, 2$ , en (3.1) define una transformación afín  $T_k$  desde la imagen fuente a la imagen destino, que puede calcularse resolviendo el sistema de ecuaciones que transforma las coordenadas de cada vértice del triángulo, esto es:

$$\boxed{p'_j = T_k \cdot p_i} \quad i = 0, 1, 2; \quad j = (k + i) \bmod 3. \quad (3.2)$$

donde  $p_i = (x_i, y_i)$  y  $p'_j = (x'_j, y'_j)$  son las coordenadas de los vértices origen y destino, respectivamente, y  $T_k$  está caracterizada por la matriz siguiente:

$$T_k = \begin{pmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.3)$$

Resolviendo (3.2) para cada una de las posibles asignaciones de vértices en (3.1), se obtienen tres transformaciones,  $T_0$ ,  $T_1$  y  $T_2$ .

En principio, todas las transformaciones resultantes de los casamientos entre triángulos pueden ser consideradas como candidatas para el registro de las imágenes de entrada. Sin embargo, en la mayoría de las aplicaciones reales, pueden establecerse límites para los valores de los parámetros de registro. Estos límites pueden emplearse para descartar muchas de las transformaciones sin necesidad de evaluarlas completamente, y reducir en consecuencia el tiempo de cómputo necesario. Los parámetros  $m_i$  empleados en (3.3), no obstante, no se corresponden de forma directa con magnitudes geométricas simples, y es conveniente emplear otra parametrización para el establecimiento de los límites.

En consecuencia, para la representación de los parámetros de una transformación afín, que tenga en cuenta traslación, rotación, escalado y sesgo, hemos considerado la estructura siguiente:

$$\begin{aligned} T &= \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & s_k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} s_x \cos \alpha & s_y (s_k \cos \alpha - \sin \alpha) & t_x \\ s_x \sin \alpha & s_y (s_k \sin \alpha + \cos \alpha) & t_y \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (3.4)$$

donde  $t_x$  y  $t_y$  indican traslación,  $s_x$  y  $s_y$  escalado,  $s_k$  sesgo y  $\alpha$  el ángulo de rotación. Como se demuestra en el apéndice A, este modelo de transformación afín es equivalente a la descomposición en valores singulares, y constituye, por tanto, una representación completa para cualquier matriz con la forma de (3.3). Para su utilización, no obstante, es necesario obtener fórmulas de conversión entre las dos parametrizaciones.

Igualando las expresiones (3.3) y (3.4) se obtienen las siguientes relaciones:

$$\boxed{t_x = m_2} \quad \boxed{t_y = m_5} \quad \boxed{\tan \alpha = \frac{m_3}{m_0}} \quad \boxed{s_x = \sqrt{m_0^2 + m_3^2}} \quad (3.5a)$$

Para obtener  $s_y$  y  $s_k$  aplicamos una rotación de  $-\alpha$  por la izquierda a (3.3) y (3.4), de donde:

$$s_y = -m_1 \sin \alpha + m_4 \cos \alpha \quad s_k = \frac{m_1 \cos \alpha + m_4 \sin \alpha}{s_y}$$

que, teniendo en cuenta las relaciones:

$$\cos \alpha = \frac{m_0}{s_x} = \frac{m_0}{\sqrt{m_0^2 + m_3^2}} \quad \sin \alpha = \frac{m_3}{s_x} = \frac{m_3}{\sqrt{m_0^2 + m_3^2}}$$

dan lugar a las expresiones siguientes:

$$\boxed{s_y = \frac{m_0 \cdot m_4 - m_1 \cdot m_3}{\sqrt{m_0^2 + m_3^2}}} \quad \boxed{s_k = \frac{m_0 \cdot m_1 + m_3 \cdot m_4}{m_0 \cdot m_4 - m_1 \cdot m_3}} \quad (3.5b)$$

Así, mediante las fórmulas (3.5a) y (3.5b), pueden calcularse los parámetros de registro correspondientes a una transformación de la forma (3.3), sobre los que se pueden aplicar de forma sencilla los límites prefijados, y descartar aquellas transformaciones que caigan fuera de ellos.

En general, los valores límite para  $t_x$ ,  $t_y$ ,  $s_x$ ,  $s_y$ ,  $s_k$  y  $\alpha$  dependerán del tipo de imágenes a registrar. Si no se dispone de ninguna información a priori sobre la relación geométrica existente entre las imágenes de entrada, puede no especificarse ningún límite para los parámetros de registro, con lo que todas las transformaciones candidatas serán evaluadas.

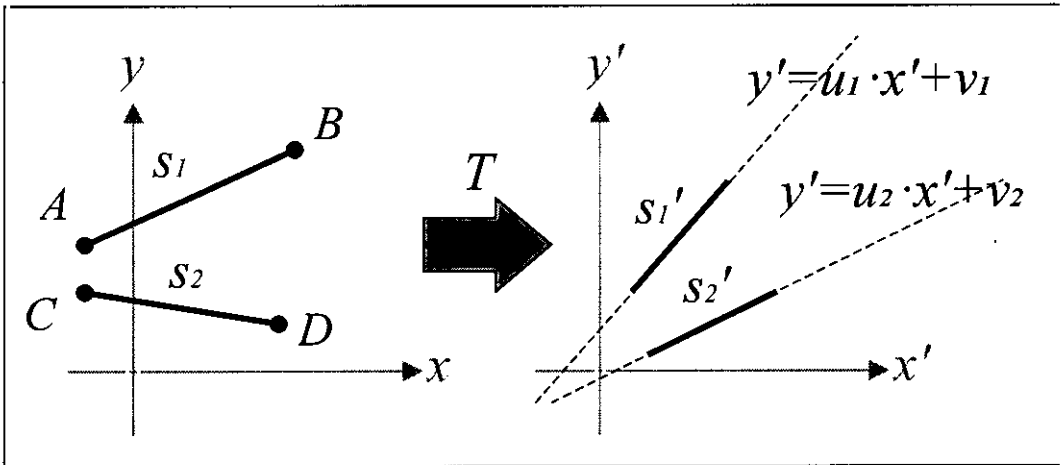
Para la prueba de registro del capítulo 5, por ejemplo, se han utilizado los siguientes límites:

$$\begin{aligned} 0.7 < s_x, s_y < 1.3 & \quad -0.2 \text{ rad} < \alpha < 0.2 \text{ rad} \\ -0.15 < s_k < 0.15 & \quad -64.0 \text{ píxeles} < t_x, t_y < 64.0 \text{ píxeles} \end{aligned} \quad (3.6)$$

### 3.3.2. Casamiento de parejas de segmentos lineales

En lugar de utilizar triángulos para la determinación de potenciales transformaciones de registro, también puede emplearse otro tipo de estructuras, como parejas de segmentos o segmentos aislados. El casamiento de estas estructuras, sin embargo, no es suficiente para determinar una transformación local afín completa. Aun así, es posible aplicar las técnicas de acumulación, que se describirán en el capítulo 4, para calcular la mejor transformación global de registro a partir de casamientos entre ellas.

En el casamiento de dos parejas de segmentos, únicamente pueden determinarse las posiciones relativas de los puntos extremos de cada uno de los segmentos fuente sobre las rectas definidas por los segmentos destino. Como consecuencia, existirán múltiples configuraciones posibles de parámetros compatibles con el casamiento de las dos parejas de segmentos.



**Figura 3.10.** Casamiento de dos parejas de segmentos,  $\{s_1, s_2\}$  con  $\{s_1', s_2'\}$ . Los vértices  $A$  y  $B$  de  $s_1$  deben caer sobre la recta definida por  $s_1'$  una vez que se aplique la transformación de registro  $T$ . Análogamente, los vértices  $C$  y  $D$  de  $s_2$  caerán sobre la recta definida por  $s_2'$ .

Supóngase la situación de la figura 3.10, en la que se determina un casamiento entre la pareja de segmentos fuente  $\{s_1, s_2\}$  y la pareja de segmentos destino  $\{s_1', s_2'\}$ . Bajo nuestro

supuesto de que solamente puede asegurarse la coincidencia de las rectas definidas por los segmentos en el proceso de registro, la única restricción que puede imponerse es que los vértices de los segmentos de la pareja fuente caigan sobre las rectas definidas por la pareja de segmentos destino. Estas rectas son:

$$\begin{aligned} s_1': \quad y' &= u_1 x' + v_1 \\ s_2': \quad y' &= u_2 x' + v_2 \end{aligned} \quad (3.7)$$

donde  $v_1, v_2$  son las coordenadas  $y$  de puntos cualesquiera de sus rectas respectivas (normalmente las de uno de los vértices de los segmentos destino) y  $u_1, u_2$  valen:

$$u_1 = \frac{s'_{1y}}{s'_{1x}} \quad u_2 = \frac{s'_{2y}}{s'_{2x}}$$

siendo  $s'_{1x}, s'_{1y}$  las componentes del vector  $s_1'$ , y  $s'_{2x}, s'_{2y}$  las del vector  $s_2'$ .

Si el segmento  $s_1$  tiene por vértices a los puntos  $A=(x_A, y_A)$  y  $B=(x_B, y_B)$ , y el segmento  $s_2$  a  $C=(x_C, y_C)$  y  $D=(x_D, y_D)$ , entonces la aplicación de una transformación de registro  $T$ , resultará en unos nuevos puntos destino  $A'=(x'_A, y'_A)$ ,  $B'=(x'_B, y'_B)$ ,  $C'=(x'_C, y'_C)$  y  $D'=(x'_D, y'_D)$ :

$$\begin{cases} (x_A & y_A & 1) \cdot T^T = (x'_A & y'_A & 1) \\ (x_B & y_B & 1) \cdot T^T = (x'_B & y'_B & 1) \\ (x_C & y_C & 1) \cdot T^T = (x'_C & y'_C & 1) \\ (x_D & y_D & 1) \cdot T^T = (x'_D & y'_D & 1) \end{cases} \quad (3.8)$$

Y, puesto que los puntos  $A'$  y  $B'$  deben caer sobre la recta definida por  $s_1'$ , y  $C'$  y  $D'$  sobre la que define  $s_2'$ , las ecuaciones (3.8) se transforman en:

$$\begin{cases} (x_A & y_A & 1) \cdot T^T = (x'_A & u_1 x'_A + v_1 & 1) \\ (x_B & y_B & 1) \cdot T^T = (x'_B & u_1 x'_B + v_1 & 1) \\ (x_C & y_C & 1) \cdot T^T = (x'_C & u_2 x'_C + v_2 & 1) \\ (x_D & y_D & 1) \cdot T^T = (x'_D & u_2 x'_D + v_2 & 1) \end{cases} \quad (3.9)$$

Entonces, si  $T$  es de la forma (3.3), operando en (3.9) se derivan las siguientes expresiones:

$$\begin{cases} m_0 x_A + m_1 y_A + m_2 = x'_A \\ m_0 x_B + m_1 y_B + m_2 = x'_B \\ m_0 x_C + m_1 y_C + m_2 = x'_C \\ m_0 x_D + m_1 y_D + m_2 = x'_D \end{cases} \quad (3.10a)$$

$$\begin{cases} m_3 x_A + m_4 y_A + m_5 = u_1 x'_A + v_1 \\ m_3 x_B + m_4 y_B + m_5 = u_1 x'_B + v_1 \\ m_3 x_C + m_4 y_C + m_5 = u_2 x'_C + v_2 \\ m_3 x_D + m_4 y_D + m_5 = u_2 x'_D + v_2 \end{cases} \quad (3.10b)$$

Sustituyendo las fórmulas de (3.10a) en (3.10b) se llega, finalmente, al siguiente sistema de cuatro ecuaciones y seis incógnitas:

$$\begin{cases} m_0 u_1 x_A + m_1 u_1 y_A + m_2 u_1 - m_3 x_A - m_4 y_A - m_5 = -v_1 \\ m_0 u_1 x_B + m_1 u_1 y_B + m_2 u_1 - m_3 x_B - m_4 y_B - m_5 = -v_1 \\ m_0 u_2 x_C + m_1 u_2 y_C + m_2 u_2 - m_3 x_C - m_4 y_C - m_5 = -v_2 \\ m_0 u_2 x_D + m_1 u_2 y_D + m_2 u_2 - m_3 x_D - m_4 y_D - m_5 = -v_2 \end{cases} \quad (3.11)$$

Puesto que es necesario determinar los seis parámetros  $m_i$  de la transformación  $T$  y tan sólo pueden plantearse cuatro ecuaciones, el sistema (3.11) no queda completamente determinado. Se obtienen, por tanto, restricciones en la forma que puede tomar  $T$ , pero para su determinación completa se necesita algún tipo de información adicional.

Cuando se trabaja con imágenes de bandas espectrales similares, puede suponerse la conservación de las posiciones relativas de los vértices de los segmentos o sus longitudes [29, 33-41], lo cual aporta información suficiente para la determinación de la transformación completa a partir de casamientos entre parejas de segmentos. En el caso más general de las imágenes de bandas espectrales diferentes no es posible realizar estas suposiciones y, para la reconstrucción de la transformación afín completa, es necesario emplear otros métodos, como los basados en técnicas de acumulación que se describen en el apartado 4.5 de esta memoria.

### 3.3.3. Casamiento de segmentos lineales aislados

A partir del casamiento de un único segmento de la imagen fuente con otro de la imagen destino el número de ecuaciones disponibles se reduce a dos, de las seis necesarias. Al



igual que ocurre con las parejas de segmentos, debe recurrirse a métodos como el descrito en la sección 4.5 para poder utilizar este tipo de casamientos.

En el casamiento de un único segmento debe tenerse también en cuenta que aparece una fase no determinada en el giro de  $\pi$  radianes, puesto que siempre existirán dos formas de asignar el segmento fuente al destino, correspondientes a los giros de ángulos  $\alpha$  y  $\alpha + \pi$ . Generalmente, debido a las limitaciones impuestas a los parámetros de registro, una de las transformaciones caerá fuera de rango, pero ambas deben ser tenidas en cuenta como posibles candidatas.

### 3.4. Evaluación de las transformaciones de registro

El proceso de casamiento de grupos de segmentos determina una lista de posibles transformaciones de registro. El paso siguiente es evaluarlas, para identificar cuál de las transformaciones candidatas tiene mayor probabilidad de registrar globalmente las imágenes de entrada. Pueden emplearse dos enfoques principales.

El primero se basa en la acumulación de los puntos asociados a cada transformación en el espacio de parámetros, con la idea de que la transformación que aparezca más frecuentemente será la que tenga mayor probabilidad de registrar las imágenes (como se propone en la referencia [51]). Este enfoque se discute en el capítulo 4.

El segundo método consiste en evaluar cada una de las transformaciones para encontrar aquella que maximice un determinado criterio de calidad. El problema en este caso es la definición de un criterio de calidad adecuado. En el registro de imágenes de bandas similares o pares estereoscópicos pueden utilizarse medidas basadas en correlación o superposición de bordes, que no son válidas en el registro de imágenes de bandas espectrales diferentes. Por ello ha sido necesario desarrollar un nuevo criterio de calidad, que se fundamenta en la idea de que, en el proceso de casamiento de segmentos en imágenes diferentes, únicamente se conserva la recta en la que está contenido cada segmento. En el capítulo 5 se estudia este método de registro.

## 4. Registro Automático basado en Técnicas de Acumulación

Para determinar la transformación global de registro a partir del casamiento local de estructuras formadas por segmentos, se han desarrollado en este trabajo dos métodos distintos. El primero de ellos, descrito en este capítulo, se fundamenta en las técnicas de acumulación utilizadas en la transformada de Hough para determinar la transformación global de registro más probable. El otro método, basado en la definición de un criterio de calidad de casamiento entre segmentos para evaluar de forma global cada una de las transformaciones candidatas, se estudia en el capítulo siguiente.

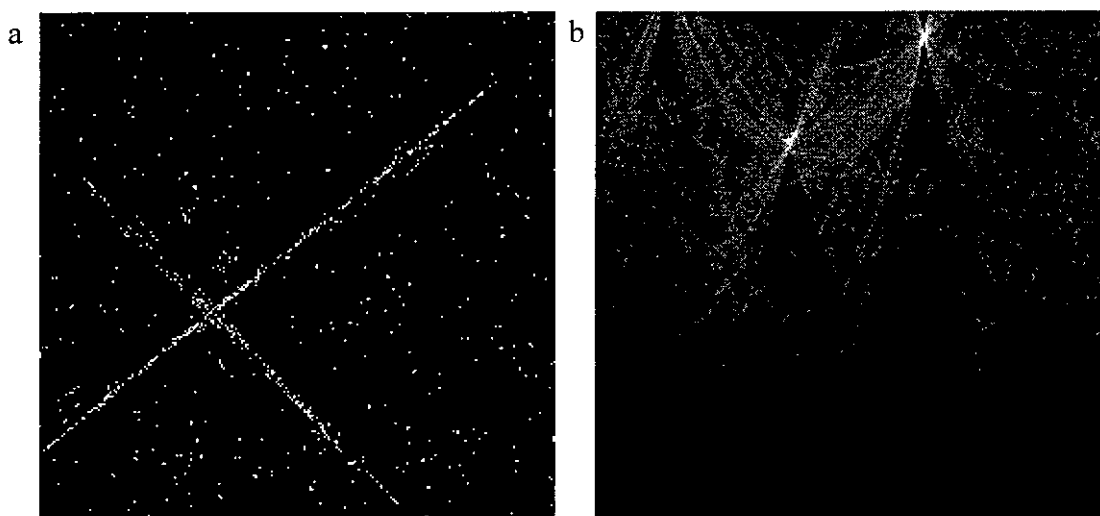
### 4.1. Introducción. Transformada de Hough

La transformada de Hough [9-13, 53-56] se emplea para la detección de rectas en imágenes binarias. Cuando se realiza una detección de bordes, es bastante común que aparezcan fragmentados. La transformada de Hough permite detectar rectas aun cuando los puntos que las componen no están todos conectados consecutivamente, e incluso cuando están muy dispersos. Posteriormente se ha extendido su uso para la detección de otras curvas [53] y, en general, para la determinación de los valores más probables para los parámetros de una función en base a un conjunto de medidas de sus resultados. A modo ilustrativo, se muestra a continuación cómo opera el mecanismo de acumulación en el método de la transformada de Hough para detectar rectas.

En el caso de la detección de rectas en una imagen binaria, cada punto de la imagen puede pertenecer a todas las rectas que pasen por él, pero solamente aquellas rectas que también pasen por otros puntos de la imagen que estén alineados con éste serán realmente líneas presentes en la imagen. Por tanto, para determinar las rectas existentes, se crea una matriz de acumulación bidimensional, asociada al espacio de acumulación de coordenadas  $(a, b)$ , en la que cada celdilla  $(a_i, b_i)$  representa una recta,  $y = a_i x + b_i$ . Para cada punto de la imagen  $(x_j, y_j)$  se incrementa en 1 el valor de las celdillas  $(a_i, b_i)$  cuya recta asociada pase por  $(x_j, y_j)$ . De este modo, una vez tenidos en cuenta todos los puntos de la imagen binaria, cada celdilla de la matriz de acumulación indicará por cuántos puntos de la imagen pasa la recta que representa. Las celdillas con mayor valor serán las que más probablemente se correspondan con rectas de la imagen. Si se desea detectar rectas de una longitud mínima determinada, puede ponerse un umbral a partir del cual se decida si existe o no recta para cada celdilla de la matriz.

En la figura 4.1 se muestra una imagen binaria con puntos dispersos en la que se aprecia una mayor densidad de puntos a lo largo de dos rectas, y su transformada de Hough, en la que destacan dos máximos, correspondientes a los parámetros de las rectas formadas por las agregaciones de puntos de la imagen binaria.

El ejemplo anterior se cita únicamente para mostrar cómo operan las técnicas de acumulación, y se ha simplificado en gran medida para mayor claridad. En la implementación real de la transformada de Hough para la detección de rectas o segmentos en una imagen binaria, no se utiliza la parametrización  $y = a \cdot x + b$ , ya que presenta singularidades en  $a$ . Se elige en su lugar una parametrización que define cada recta mediante su dirección y su distancia mínima al origen [9-13], y es la que ha sido empleada para la generación de la imagen 4.1b. Además, para decidir si el segmento o recta está presente en la imagen, la condición de que el número de puntos acumulados en cada celdilla supere un cierto umbral es condición necesaria pero no suficiente para que constituyan una recta o segmento, ya que no se tiene en cuenta si los píxeles acumulados están o no conectados, lo cual es de gran importancia para realizar la decisión (y más aún si las longitudes que se pretende detectar son pequeñas). Una forma de mejorar el criterio de decisión podría consistir en dar un peso diferente a cada punto, en lugar de acumular siempre el valor de 1, en función de la disposición de los píxeles de su vecindad 3x3, por ejemplo. De ese modo se modifica la función de peso utilizada, pero la técnica de acumulación sigue siendo la misma.



**Figura 4.1.** (a) Imagen binaria, en la que se aprecian agregaciones de puntos a lo largo de dos rectas, y (b) su transformada de Hough.

## 4.2. Técnicas de acumulación aplicadas al registro de imágenes

La idea básica del método de acumulación aplicado al registro de imágenes consiste en determinar la mejor transformación global a partir de la evaluación de transformaciones locales, contabilizando la frecuencia de aparición de cada una de ellas. Las transformaciones locales pueden determinarse por múltiples métodos, como por ejemplo, utilizando casamientos entre triángulos fuente y destino. Nuestro método de registro por acumulación está basado en la suposición de que la mejor transformación de registro se corresponde con la que más elementos (triángulos, en el ejemplo) es capaz de casar entre fuente y destino. Por lo tanto, esa transformación debe ser la que aparezca en más ocasiones al efectuar los casamientos locales.

Para la determinación de la transformación global óptima, se crea una matriz (la matriz de acumulación) en la que se van contabilizando las ocurrencias de las transformaciones asociadas a cada casamiento entre un elemento fuente y otro destino. Al final, la celdilla de la matriz que contenga un mayor número de ocurrencias será la que previsiblemente se corresponda con los parámetros de la mejor transformación de registro. En nuestro método no se contabilizan únicamente las apariciones, sino un valor relacionado con el peso o calidad de cada transformación local. Ese peso puede determinarse de múltiples formas, como, por ejemplo, calculando cuánto se superponen los segmentos que se están casando como resultado de la aplicación de la transformación local en cuestión.

A continuación se describen en líneas generales las técnicas de acumulación. Posteriormente se estudia el caso concreto de su aplicación al registro de imágenes.

### 4.2.1. Descripción general del método de acumulación

En general, el método de acumulación para la determinación de los valores más probables para el conjunto de  $n$  parámetros

$$P = \{p_1, \dots, p_n\},$$

que caracteriza una función

$$F(P) \quad \text{tal que} \quad y = F(P)(x),$$

a partir de un conjunto de  $m$  medidas

$$X = \{(x_1, y_1), \dots, (x_m, y_m)\},$$

se basa en la creación de un espacio de acumulación de dimensión  $n$  en el que se van contabilizando las calidades o pesos,  $Q(P^k)$ , de todos los conjuntos de parámetros  $P^k$  compatibles con cada medida realizada  $(x_j, y_j)$ .

En la mayoría de las ocasiones únicamente se considera una región,  $\mathfrak{R}$ , del espacio de acumulación, bien porque ya ha sido acotado previamente el rango de valores posible o debido a limitaciones en el espacio de almacenamiento. Dicha región está comprendida entre dos conjuntos de parámetros (vectores del espacio de acumulación, en definitiva) máximo y mínimo,  $P^{min} = \{p_1^{min}, \dots, p_n^{min}\}$  y  $P^{max} = \{p_1^{max}, \dots, p_n^{max}\}$ , esto es:

$$p_i^{min} \leq p_i < p_i^{max} \quad \forall P = \{p_i\} \in \mathfrak{R}; \quad i = 1, \dots, n$$

La precisión en la determinación de los parámetros está limitada por un conjunto de valores  $dP = \{dp_1, \dots, dp_n\}$ , que depende del método utilizado para estimarlos. De este modo, el espacio de acumulación queda dividido en  $(c_1^{max}+1) \cdot (c_2^{max}+1) \cdot \dots \cdot (c_n^{max}+1)$  celdillas que conforman una matriz,  $M(C)$ , a la que se denomina matriz de acumulación.  $C = \{c_1, \dots, c_n\}$  es el vector de indexación de la matriz, para el que cada uno de los índices  $c_i$  toma los valores:

$$0 \leq c_i \leq c_i^{max} = \left\lceil \frac{p_i^{max} - p_i^{min}}{dp_i} \right\rceil \quad c_i \in \mathbf{Z}^+ \quad (4.1)$$

donde los corchetes indican parte entera.

Al discretizar el espacio de acumulación según la precisión  $dP$ , los pesos de todos los conjuntos de parámetros  $P = \{p_i\}$  que cumplan:

$$k_i \cdot dp_i \leq p_i < (k_i + 1) \cdot dp_i; \quad 0 \leq k_i \leq c_i^{max} \in \mathbf{Z}^+$$

se acumularán en la misma celdilla,  $C(P)$ :

$$C(P) = \{c_1(P), c_2(P), \dots, c_n(P)\} \quad (4.2)$$

donde cada índice  $c_i$  vale:

$$c_i(P) = c_i(p_i) = \left\lceil \frac{p_i - p_i^{\min}}{dp_i} \right\rceil \quad (4.3)$$

Una vez realizadas todas las acumulaciones asociadas al conjunto de medidas  $X$  el valor de cada celdilla  $C$  de la matriz de acumulación será el siguiente:

$$M(C) = \sum_{j=1}^m Q(P = P(C) \mid y_j = F(P)(x_j)) \quad (4.4)$$

donde cada conjunto de parámetros  $P(C)$  es:

$$P(C) = \{p_i(C)\} = \{p_i(c_i)\} = \{c_i \cdot dp_i + p_i^{\min}\}; \quad i = 1, \dots, n \quad (4.5)$$

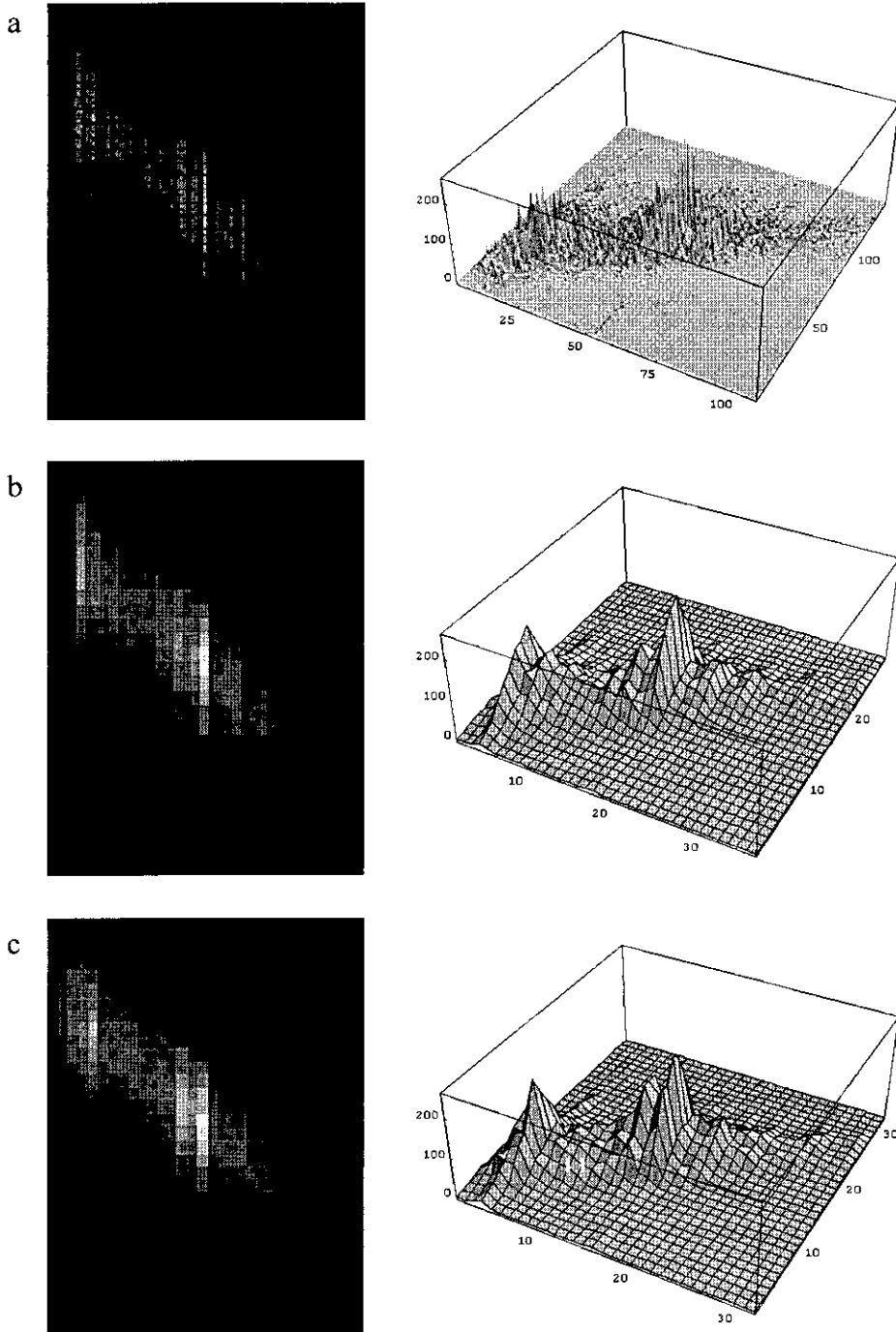
Después se determina la celdilla  $C^0$  con peso acumulado máximo, que se corresponderá con el mejor conjunto de parámetros  $P(C^0)$  para la función  $F$ , compatible con el conjunto de medidas  $X$ . Dependiendo de la aplicación para la que se calculen los parámetros de la función  $F$ , puede ser interesante extraer no sólo el máximo de acumulación, sino todos aquellos conjuntos de parámetros  $P^k$  para los cuales se sobrepase un umbral determinado. Es el caso, por ejemplo, de la detección de rectas en imágenes binarias, donde se umbraliza a partir de un número determinado de puntos.

#### 4.2.2. Determinación de las precisiones de acumulación

La utilización de una estructura de acumulación discreta pone un límite a las precisiones de los parámetros que deben ser evaluados. Estas precisiones impondrán límites a la exactitud de los resultados del método de registro basado en la técnica de acumulación. Por esta razón, es necesario determinar la precisión  $dp_i$  con la que puede ser estimado cada parámetro de la transformación de registro.

El cálculo de las precisiones de acumulación es, además, imprescindible para garantizar el funcionamiento de la técnica de acumulación. Como puede verse en la figura 4.2, la utilización de una precisión mayor que la alcanzable por el algoritmo crea dispersión en los puntos acumulados, con lo que el máximo real puede no ser apreciable. Precisiones menores que la alcanzable pueden crear máximos espurios, al agrupar celdillas que, a la

precisión adecuada, estarían separadas y tendrían menor valor que el máximo real, pero cuya suma de valores supera el valor del mismo.



**Figura 4.2.** Diferencia en los resultados de la acumulación (para un plano ángulo-escalado) en función de las precisiones de acumulación utilizadas: (a) acumulación con precisión superior a la requerida (tamaño de la matriz 2D, 108x144), que provoca dispersión, (b) precisión correcta (tamaño 36x29), (c) precisión ligeramente diferente a la necesaria (tamaño 32x32), que refuerza algunos máximos secundarios.

Por otra parte, las precisiones  $dp_i$ , y los límites impuestos para los valores de cada parámetro,  $p_i^{min}$  y  $p_i^{max}$ , determinan el tamaño del espacio de almacenamiento necesario para la matriz de acumulación. En ese sentido, su cálculo es de gran importancia a la hora de realizar una implementación práctica de los métodos de acumulación, ya que los recursos de memoria son siempre limitados. Con frecuencia, son los límites en la memoria de almacenamiento los que imponen restricciones sobre los límites y precisión de los parámetros de la función, ya que el número de celdillas de la matriz crece de forma exponencial con el número de parámetros que deben ser calculados.

En el caso particular del registro, la precisión está directamente relacionada con el proceso de casamiento de los segmentos. Allí, la discretización en la posición de sus vértices, determinada por la utilización de coordenadas  $(x, y)$  enteras en el espacio imagen, impone límites a las variaciones mínimas medibles para los parámetros de la transformación de registro. A continuación se calculará cómo esa discretización del espacio imagen, en unidades de píxel, limita las precisiones de cada uno de los parámetros de una transformación afín de registro.

Para cada uno de los parámetros de la transformación de registro, su mínima variación discernible vendrá dada por el cambio de su valor asociado a la variación mínima (de un píxel) en la posición de uno de los vértices de un segmento.

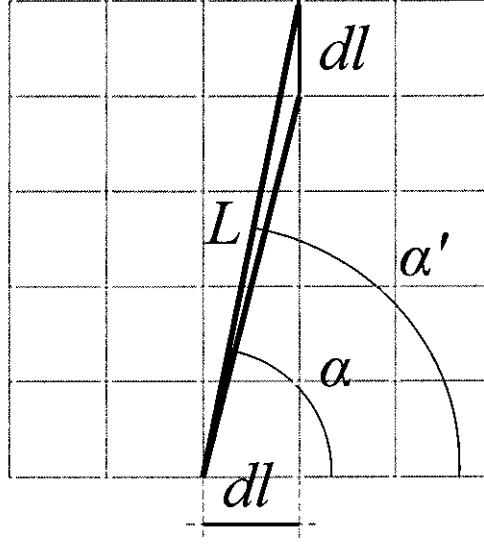
En el caso de la traslación, el límite de precisión es directamente la unidad del espacio imagen,  $dl$ , es decir:

$$dt_x = dt_y = dl = 1 \text{ píxel} \quad (4.6)$$

Para los demás parámetros de la transformación de registro, las precisiones están directamente relacionadas con la longitud de los segmentos que se utilicen para calcularlos. En este sentido, debe especificarse una longitud,  $L$ , que se corresponda con la longitud media de los segmentos detectados. Puesto que esa longitud no puede ser superior a la de una de las diagonales de la imagen, puede tomarse la longitud de la diagonal como valor para  $L$ . También puede utilizarse la longitud del mayor segmento detectado. En general sólo estamos interesados en el orden de magnitud mínimo de las precisiones asociadas a la discretización del espacio imagen, ya que posteriormente deberán ser reajustadas dependiendo de la precisión del método de registro elegido y de la cantidad de memoria disponible para realizar las acumulaciones. Se utilizará, entonces,  $L$  como parámetro libre en las fórmulas de la precisión de los demás parámetros, y se ajustará su valor para la situación concreta que se esté tratando.



En el caso de la estimación del giro, la variación mínima discernible asociada a un desplazamiento de 1 píxel se corresponde con el diagrama de la figura 4.3.



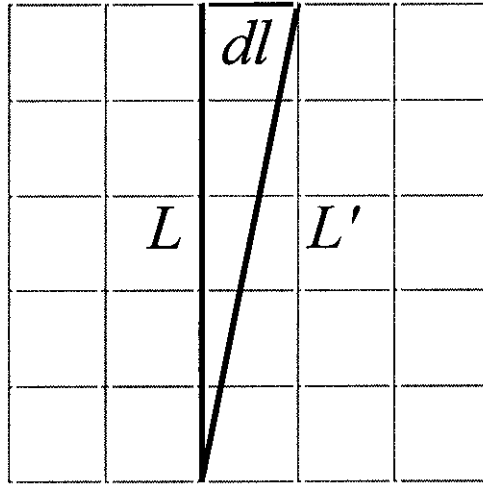
**Figura 4.3.** Mínima variación detectable en giro mediante la variación de la posición de un vértice equivalente a una longitud  $dl$  a lo largo de uno de los ejes de coordenadas.

El valor de  $d\alpha$  será, entonces:

$$\begin{aligned}
 \alpha &= \tan^{-1}\left(\frac{y}{x}\right); \\
 d\alpha &= \alpha' - \alpha = \tan^{-1}\left(\frac{y'}{x}\right) - \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{\sqrt{L^2 - dl^2}}{dl}\right) - \tan^{-1}\left(\frac{\sqrt{L^2 - dl^2} - dl}{dl}\right) = \\
 &= \tan^{-1}\left(\sqrt{\frac{L^2}{dl^2} - 1}\right) - \tan^{-1}\left(-1 + \sqrt{\frac{L^2}{dl^2} - 1}\right) \quad (4.7)
 \end{aligned}$$

Para el escalado, la situación de mínima variación se muestra en la figura 4.4. En ese caso su precisión,  $ds$ , es:

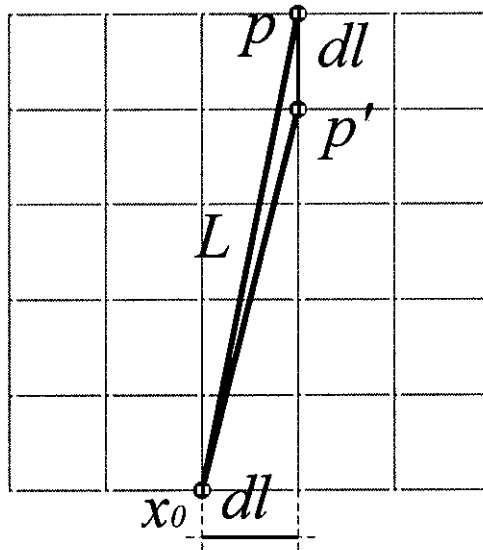
$$\begin{aligned}
 s &= \frac{l_2}{l_1}; \\
 ds &= s' - s = \frac{L'}{L} - \frac{L}{L} = \frac{\sqrt{L^2 + dl^2}}{L} - 1 = \sqrt{1 + \left(\frac{dl}{L}\right)^2} - 1 \quad (4.8)
 \end{aligned}$$



**Figura 4.4.** Mínima variación detectable en escalado mediante la variación de la posición de un vértice equivalente a una longitud  $dl$  a lo largo de uno de los ejes de coordenadas.

El esquema para la determinación de la precisión del sesgo puede verse en la figura 4.5. El valor de  $ds_k$  queda:

$$\begin{aligned}
 s_k &= \frac{x - x_0}{y} ; \\
 ds_k &= s'_k - s_k = \frac{x' - x_0}{y'} - \frac{x - x_0}{y} = \frac{dl}{\sqrt{L^2 + dl^2}} - \frac{dl}{\sqrt{L^2 + dl^2}} = \\
 &= \frac{dl^2}{L^2 + dl^2 - dl\sqrt{L^2 + dl^2}}
 \end{aligned} \tag{4.9}$$



**Figura 4.5.** Mínima variación detectable en sesgo mediante la variación de la posición de un vértice equivalente a una longitud  $dl$  a lo largo de uno de los ejes de coordenadas.

La elección de un valor para  $L$  debe basarse en la estimación de una longitud promedio adecuada. A tal fin se han extraído los histogramas de las longitudes de los segmentos detectados en cuatro imágenes reales de muestra, y se ha calculado el valor medio de las longitudes de los segmentos. Se han empleado los parámetros que habitualmente utilizamos para las pruebas de registro visible-IR:

tamaño de las imágenes	320 x 240 píxeles
longitud mínima de los segmentos considerados	16 píxeles
anchura de la gaussiana para el extractor de bordes	1.5 píxeles
distancia máxima para la aproximación por segmentos	1.5 píxeles

Los histogramas, con longitudes desde 0 hasta 100, se muestran en la figura 4.6.



**Figura 4.6.** Histogramas de las longitudes (de 0 a 100 píxeles) de los segmentos detectados en cuatro imágenes de prueba.

Las longitudes medias son las siguientes:

$$\left. \begin{array}{l} \bar{l}_a = 27.5 \text{ píxeles} \\ \bar{l}_b = 29.3 \text{ píxeles} \\ \bar{l}_c = 27.0 \text{ píxeles} \\ \bar{l}_d = 23.6 \text{ píxeles} \end{array} \right\} \bar{l} = 26.8 \text{ píxeles} \quad (4.10)$$

Para un valor de  $L$  de 26.8 píxeles, los valores en las precisiones de acumulación son:

$$\begin{aligned} dt_x &= dt_y = dt = 1 \text{ píxel} \\ d\alpha &= 0.0015 \text{ rad} = 0.083^\circ \\ ds_x &= ds_y = ds = 7 \cdot 10^{-4} = 0.07 \% \\ ds_k &= 0.0014 \end{aligned} \quad (4.11)$$

Para un valor de  $L$  del orden del tamaño de la imagen se obtienen los límites inferiores para las precisiones de acumulación. En imágenes de 256 píxeles de lado esos límites son:

$$\begin{aligned}
dt_x &= dt_y = dt = 1 \text{ píxel} \\
d\alpha &= 1.5 \cdot 10^{-5} \text{ rad} = 8.8 \cdot 10^{-4}^\circ \\
ds_x &= ds_y = ds = 7.6 \cdot 10^{-6} = 7.6 \cdot 10^{-4} \% \\
ds_k &= 1.5 \cdot 10^{-5}
\end{aligned} \tag{4.12}$$

Y, como puede verse, las precisiones son altas para un amplio rango de valores de  $L$ , salvo en el caso de la traslación, que únicamente depende de  $dl$ .

### 4.3. Reducción de la dimensionalidad del espacio de acumulación

La cantidad de memoria necesaria,  $\Omega$ , para almacenar la matriz de acumulación, suponiendo un tamaño de dato para cada celdilla de 4 bytes (flotante de simple precisión) es la siguiente:

$$\Omega = 4 \cdot \prod_{i=1}^n \left( \left\lceil \frac{p_i^{\max} - p_i^{\min}}{dp_i} \right\rceil + 1 \right) = 4 \cdot \prod_{i=1}^n (c_i^{\max} + 1) \text{ bytes} \tag{4.13}$$

Suponiendo un mismo número de divisiones,  $c_i^{\max} + 1 = N$ , para cada uno de los ejes coordenados del espacio de acumulación, es decir, de valores posibles para cada parámetro dentro de su rango permitido de valores,  $p_i^{\min} \leq p_i < p_i^{\max}$ , el tamaño de la matriz de acumulación crece como:

$$\Omega \approx 4 \cdot N^n \tag{4.14}$$

Los requisitos de almacenamiento, por tanto, crecen exponencialmente con el número de parámetros de la transformación de registro. Así, por ejemplo, para determinar mediante acumulación una transformación de segundo orden completa (12 parámetros), suponiendo únicamente 16 divisiones por eje (insuficiente para la mayoría de las aplicaciones), ya se obtiene un tamaño de  $10^{15}$  bytes (unos 1000 Terabytes) una cantidad muy superior a la capacidad de almacenamiento de las computadoras actuales. En consecuencia, no resulta factible restaurar una transformación de registro de segundo orden completa mediante técnicas basadas en acumulación.

Para la restauración de una transformación afín se necesitan 6 parámetros. Usando de nuevo sólo 16 divisiones por eje se obtiene un tamaño de 67 Megabytes, del orden de la

memoria máxima de un ordenador moderno, lo cual convierte la determinación de una transformación afin completa en una tarea complicada.

Debe tenerse en cuenta que la utilización de cantidades de memoria grandes no sólo supone un problema de espacio, sino también de tiempo, ya que el ordenador se verá obligado a realizar numerosas operaciones de intercambio a disco (*swapping*) para acceder a la matriz de acumulación. Por otra parte, el escaso número de divisiones en las que puede ser dividido el rango de valores de cada parámetro limita enormemente la capacidad de funcionamiento de la técnica de acumulación, ya que, o bien se consideran rangos muy pequeños para los parámetros (con lo que se restringe la búsqueda a regiones muy reducidas del espacio de acumulación), o bien se acumula con muy baja precisión (lo cual genera una gran cantidad de máximos falsos, que deberán ser reevaluados con precisiones mayores para decidir sobre su calidad). La utilización de precisiones del orden de las determinadas en el apartado anterior, como sería deseable, reducen la zona del espacio de acumulación a unos límites demasiado estrechos como para que pueda realizarse algún tipo de búsqueda del máximo si no se conoce de antemano su situación aproximada. Es necesario, por tanto, emplear algún método para evitar estas dificultades.

En ese sentido, hemos desarrollado una técnica que descompone el espacio de acumulación en subespacios de dimensiones menores, mediante el colapso de uno o varios ejes del mismo. Se pierde, en consecuencia, discernibilidad en algunas coordenadas, pero se reduce enormemente la cantidad de memoria necesaria para el almacenamiento de los parámetros. Por comodidad de representación y manejo, hemos utilizado la proyección del espacio de acumulación sobre planos bidimensionales, definidos por cada dos ejes del espacio. De esta forma, un espacio de acumulación de dimensión  $n$  queda descompuesto en  $n/2$  proyecciones bidimensionales. Cada plano bidimensional  $M_{u,v}$  de acumulación de ejes  $u$  y  $v$ , contiene los siguientes valores:

$$M_{u,v}(j_u, j_v) = \sum_{i \neq u,v} \sum_{k=0}^{c_i^{max}} M(C \mid c_i = k, c_u = j_u, c_v = j_v) \quad (4.15)$$

donde  $j_u$  y  $j_v$  son las coordenadas de indexación de la nueva matriz bidimensional, que recorren los mismos valores que  $c_u$  y  $c_v$  en la matriz de acumulación  $M$ , es decir:

$$j_u = 0, \dots, c_u^{max} \quad j_v = 0, \dots, c_v^{max} \quad (4.16)$$

La resolución en cada eje puede aumentarse considerablemente mediante esta técnica. El nuevo número de divisiones,  $N'$ , para el mismo tamaño de almacenamiento,  $\Omega$ , crece como:

$$\Omega \approx 4 \cdot N^n = 4 \cdot \frac{n}{2} \cdot (N')^2 \Rightarrow N' = \sqrt{\frac{2N^n}{n}} \quad (4.17)$$

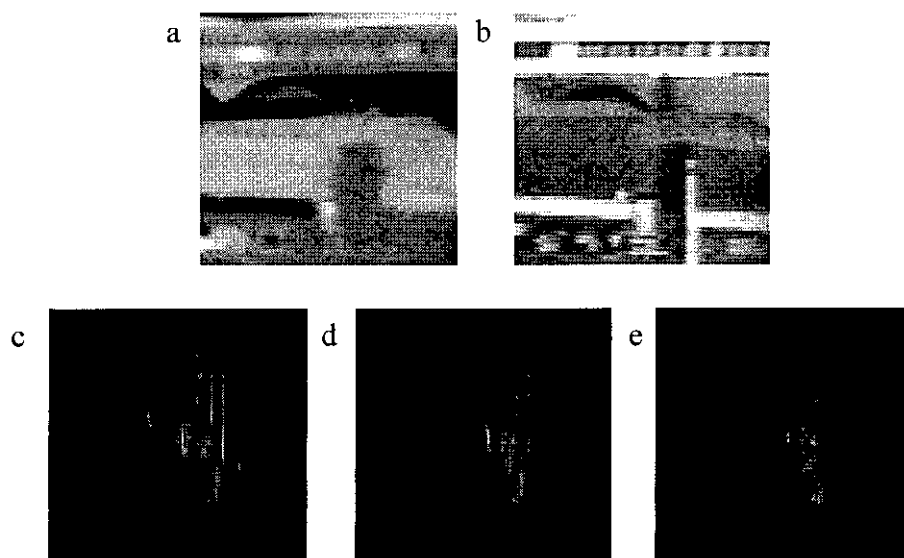
De este modo, con la misma cantidad de memoria de almacenamiento, se puede aumentar en gran medida la resolución en cada eje. Como contrapartida, la proyección del espacio de acumulación sobre uno de sus planos coordenados genera ruido de acumulación en el plano, que reduce la discernibilidad del máximo de acumulación y puede generar máximos de acumulación espurios.

Un modo de reducir este ruido de acumulación asociado a la proyección consiste en acotar las regiones proyectadas. En la figura 4.7 puede apreciarse la reducción del ruido de acumulación como consecuencia de la restricción de la región del espacio que se proyecta sobre los planos de acumulación bidimensionales. En la figura se muestra un par de imágenes IR-visible ya registradas. El máximo de acumulación en traslación, por tanto, debe aparecer en el centro del plano de acumulación  $t_x-t_y$  (correspondiente a la traslación nula). En la imagen 4.7c ese máximo principal aparece rodeado de máximos secundarios, cuya importancia disminuye al restringir la coordenada de rotación (4.7d). Después de fijar también el valor del escalado (figura 4.7e) se observa una reducción del ruido aún mayor.

Algunas coordenadas suelen ser más fáciles de determinar que otras. La coordenada de giro, en particular, suele ser fácil de determinar utilizando el llamado efecto de pico (*peaking-effect*) [57], que aparece al obtener los histogramas de las orientaciones de los segmentos detectados. Puesto que el giro se conserva bajo las transformaciones de escalado y traslación, es posible correlacionar directamente los histogramas de los conjuntos de segmentos detectados en dos imágenes de una misma escena para obtener el ángulo de rotación que las registra. La proyección del espacio de acumulación sobre el eje del parámetro de rotación es similar a la realización de esa correlación de orientaciones y puede utilizarse para restringir en una dimensión la región del espacio considerada.

Esta técnica, sin embargo, es poco robusta frente a la aparición de nuevos segmentos en una de las imágenes (debida a un escalado severo, por ejemplo) o a la existencia de deformaciones de sesgo o de órdenes superiores al lineal, y además no puede emplearse con las demás coordenadas. Por otra parte, la proyección del espacio sobre un único eje

lleva también asociado un incremento del ruido de acumulación aún más acusado que el que aparece en las proyecciones bidimensionales.



**Figura 4.7.** (a), (b) Dos ventanas de un par de imágenes IR-visible ya registradas. (c) Proyección de la región considerada del espacio de acumulación sobre el plano de traslaciones. (d) Proyección de la misma región una vez restringida la coordenada de giro. (e) Nueva proyección, después de restringir también el escalado.

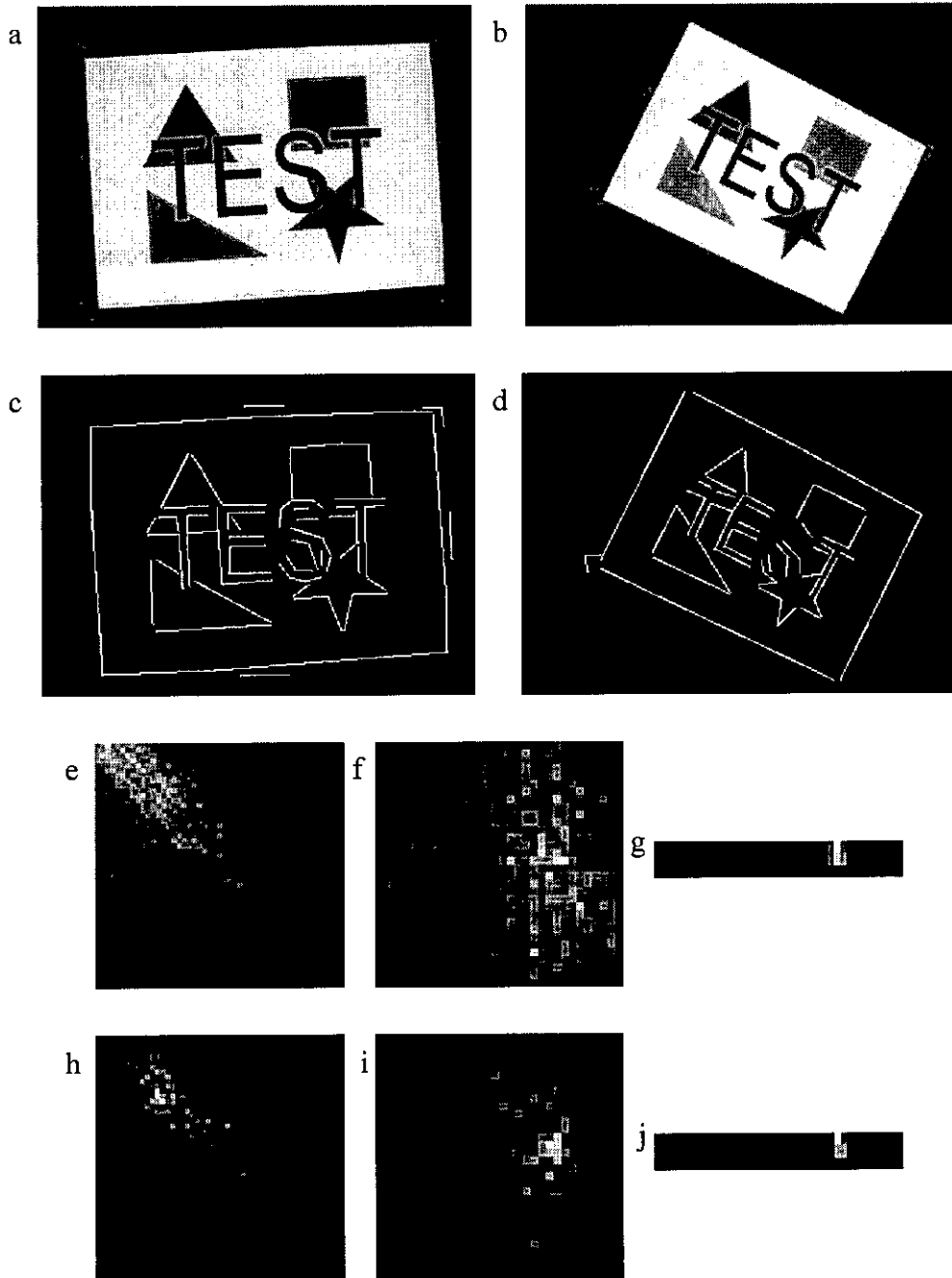
Deben buscarse, por tanto, otras formas de reducir el ruido de acumulación. Una de ellas consiste en elegir adecuadamente la función de peso que se acumula para cada transformación.

#### 4.4. Determinación de la función de peso

Cada transformación afín obtenida mediante el casamiento de dos estructuras locales debe añadirse a la matriz de acumulación en la celdilla correspondiente a los parámetros que la definen. Es importante determinar una función de peso adecuada para aumentar en lo posible la discernibilidad del máximo de acumulación, ya que la función de peso más sencilla, consistente en acumular un valor constante (típicamente 1) por cada transformación, genera con facilidad máximos falsos o distribuciones muy uniformes sobre las proyecciones bidimensionales.

La longitud de superposición de los segmentos que componen las estructuras que están siendo casadas puede utilizarse como función de peso, aunque ya se ha mencionado que no es un atributo adecuado para el registro de imágenes de bandas espectrales diferentes. No

obstante, para el registro de imágenes similares, cuando sólo es necesario restaurar un pequeño número de parámetros o las deformaciones a restaurar son muy pequeñas, los métodos de acumulación basados en superposición dan buenos resultados.



**Figura 4.8.** Diferencias en los resultados de la acumulación según la función de peso utilizada. (a), (b) Par de imágenes de la misma banda espectral. (c), (d) Segmentos detectados. Planos de (e) escalado, (f) traslaciones y (g) ángulo/sesgo para la función de peso unidad. (h), (i), (j) Planos de acumulación empleando la longitud de superposición entre los segmentos que componen los triángulos casados como función de peso.



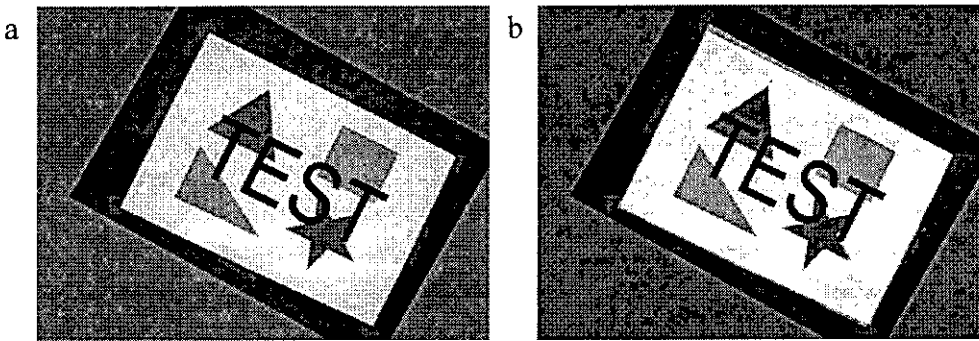
En la figura 4.8 se pone de manifiesto la clara diferencia que supone la utilización de la longitud de superposición como función de peso, frente a la acumulación de un peso constante. Aunque el ángulo se restaura correctamente con las dos funciones de peso (debido al efecto de pico), los máximos de los diagramas de acumulación para las coordenadas de traslación y escalado no se corresponden con los parámetros de la transformación de registro cuando se utiliza la función de peso unidad. La región del espacio de parámetros considerada ha sido la siguiente:

$$\begin{aligned} 0.5 &\leq s_x, s_y < 1.5 \\ -1 \text{ rad} &\leq \alpha < 1 \text{ rad} \\ -0.01 &\leq s_k < 0.01 \\ 32 \text{ píxeles} &\leq t_x, t_y < 32 \text{ píxeles} \end{aligned}$$

La transformación de registro afin obtenida, a partir de los diagramas de acumulación con la función de peso longitud (figuras 4.8h, i, j) es la siguiente:

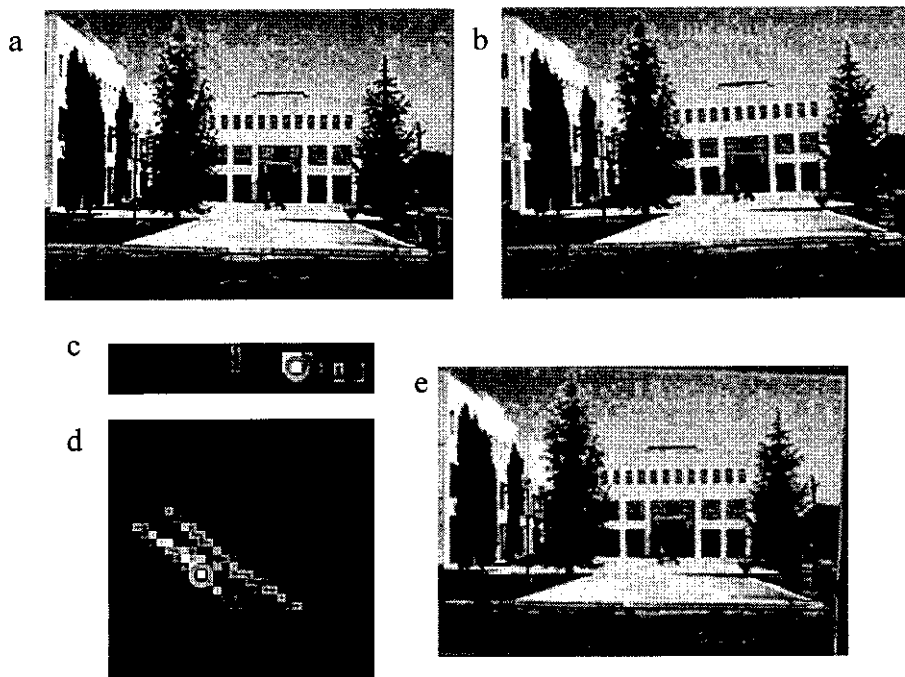
$$\begin{aligned} s_x &= 0.773658 & s_y &= 0.751265 \\ \alpha &= 0.512600 \text{ rad} & s_k &= -0.002585 \\ t_x &= 15.211648 \text{ píxeles} & t_y &= -0.884711 \text{ píxeles} \end{aligned}$$

En la figura 4.9a se muestra el resultado de la aplicación de esta transformación de registro a la imagen 4.8a. En el resultado de la fusión de las imágenes 4.9a y 4.8b, recogido en la figura 4.9b, se observa una ligera duplicación de bordes, debida a la deformación de distorsión (no lineal, de tercer orden) que presenta la óptica de la cámara utilizada.



**Figura 4.9.** (a) Registro de la imagen 4.8a con la 4.8b utilizando la transformación de registro obtenida mediante la longitud de superposición como función de peso. (b) Fusión de 4.9a y 4.8b.

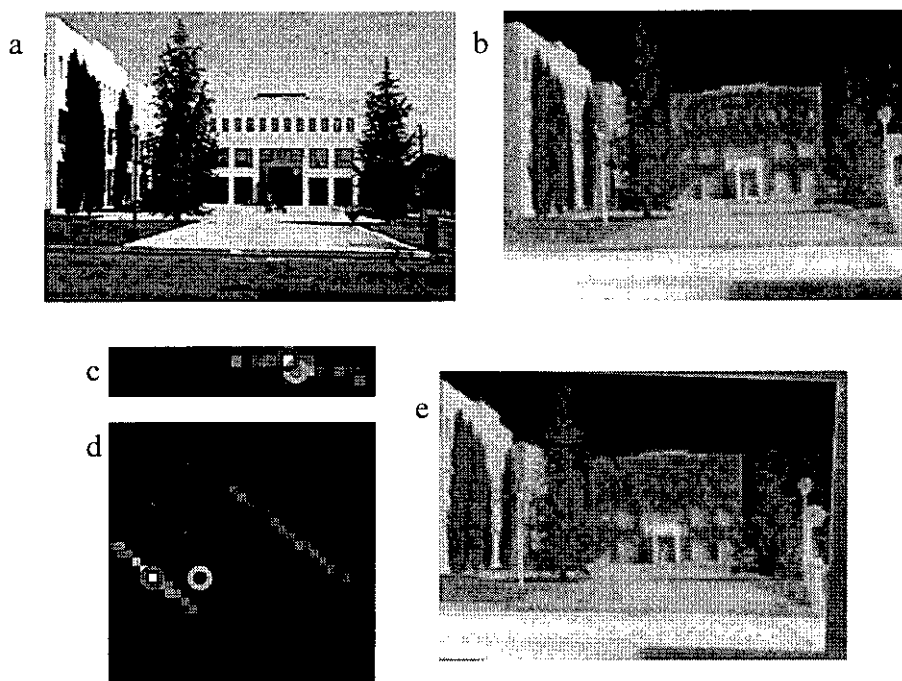
La longitud de superposición local de los segmentos, no obstante, no es una función de peso adecuada para el registro de imágenes de bandas espectrales diferentes (tal y como se discutía en la sección 3.1 sobre elementos invariantes). Este hecho se pone de manifiesto en las figuras 4.10 y 4.11, en las que se comparan los resultados obtenidos mediante el registro por acumulación de una pareja visible-visible (figura 4.10) y visible-IR (figura 4.11) desregistradas mediante la misma transformación. Como puede observarse, únicamente en el caso visible-visible se obtienen los máximos de acumulación en las posiciones correctas de sus planos (círculos claros). En el caso visible-IR los máximos obtenidos (círculos oscuros) no se corresponden con la transformación de registro correcta, que globalmente sí que puede suponer una superposición global máxima.



**Figura 4.10.** Registro de dos imágenes similares mediante acumulación con longitud de superposición como función de peso. (a) Imagen original, y (b) par de registro artificial, obtenido a partir de (a) mediante una rotación de  $2^\circ$  y una traslación de  $(-5, 2)$  píxeles. Planos de acumulación en (c) ángulo-sesgo y (d) traslaciones. (e) Imagen (b) registrada en función de los resultados de la acumulación. Las coordenadas del máximo de acumulación coinciden con las de la deformación aplicada (círculos).

La principal conclusión que puede sacarse de ejemplos como este, es que, efectivamente, no puede confiarse en la realización de medidas locales para la determinación de una transformación global óptima para el registro de imágenes de bandas espectrales alejadas, en consonancia con la discusión sobre elementos invariantes de la sección 3.1. No existe, por tanto, ninguna función de peso adecuada para el registro por acumulación de imágenes de bandas espectrales alejadas. En consecuencia, un método de registro general, para

cualquier tipo de imágenes, debe validar las transformaciones realizando medidas globales sobre los conjuntos de segmentos detectados. Eso no reduce, sin embargo, la utilidad de los métodos de acumulación, que pueden ser empleados para la resolución de variados problemas de registro. Además de su utilización para el registro de imágenes similares, pueden utilizarse también para otras tareas específicas, como el registro visible-SAR (sección 8.2) o el reconocimiento de formas.



**Figura 4.11.** Registro de dos imágenes de bandas espectrales alejadas (visible-IR) mediante acumulación con longitud de superposición como función de peso. (a) Imagen original, y (b) par de registro artificial, obtenido a partir de una imagen IR registrada con (a), mediante la aplicación de la misma transformación de desregistro utilizada en la figura 4.10. Planos de acumulación en (c) ángulo-sesgo y (d) traslaciones. (e) Imagen (b) registrada en función de los resultados de la acumulación. Los máximos de acumulación (círculos oscuros) no coinciden con los parámetros de la transformación aplicada (círculos claros).

#### 4.5. Acumulación de transformaciones afines incompletas

El casamiento de algunos tipos de estructuras locales no es suficiente para determinar una transformación afín completa. Es el caso del casamiento de parejas de segmentos (sección 3.3.2), en el que el sistema de ecuaciones asociado (3.9) no queda completamente determinado. Un método para resolver esa indeterminación consiste en aprovechar la naturaleza discreta de la matriz de acumulación. Puesto que el espacio de acumulación está discretizado, siempre es posible realizar un barrido en dos de los ejes, dando valores a sus

parámetros de registro, de modo que solamente queden cuatro parámetros libres en el sistema de ecuaciones. Una vez resuelto el sistema para cada par de parámetros fijados, se aplica la transformación obtenida a los segmentos fuente y se determina el peso correspondiente al casamiento con los segmentos destino, el cual se añade a la matriz de acumulación en la posición que corresponda a la transformación.

Esta técnica puede extenderse al casamiento de un único segmento fuente con uno destino (sección 3.3.3). En tal caso, el número de ecuaciones disponibles se reduce a dos, de las seis necesarias. De nuevo es posible barrer cuatro de las dimensiones del espacio de acumulación para obtener transformaciones que serán posteriormente evaluadas con la función de peso. El método de registro visible-SAR, descrito en la sección 8.2, emplea este tipo de casamientos.

Una de las ventajas que tiene el empleo de estructuras más simples que los triángulos es que su número es menor, con lo que el número total de casamientos es también menor, y se reducen el tiempo y la capacidad de cómputo necesarios. En efecto, suponiendo que el número de segmentos del conjunto fuente es  $N_1$  y el del conjunto destino es  $N_2$ , y que ambos son aproximadamente del mismo orden de magnitud,  $N$ :

$$N_1 \approx N_2 \approx N$$

entonces, el número total de casamientos,  $N_c$ , que deben ser evaluados es:

$$N_c = C_{N_1,3} \cdot C_{N_2,3} \approx (C_{N,3})^2 = \left( \frac{N!}{3!(N-3)!} \right)^2 = \left( \frac{N \cdot (N-1) \cdot (N-2)}{6} \right)^2 \approx N^6$$

Mientras que para el caso de casamientos entre parejas de segmentos esa cantidad se reduce a:

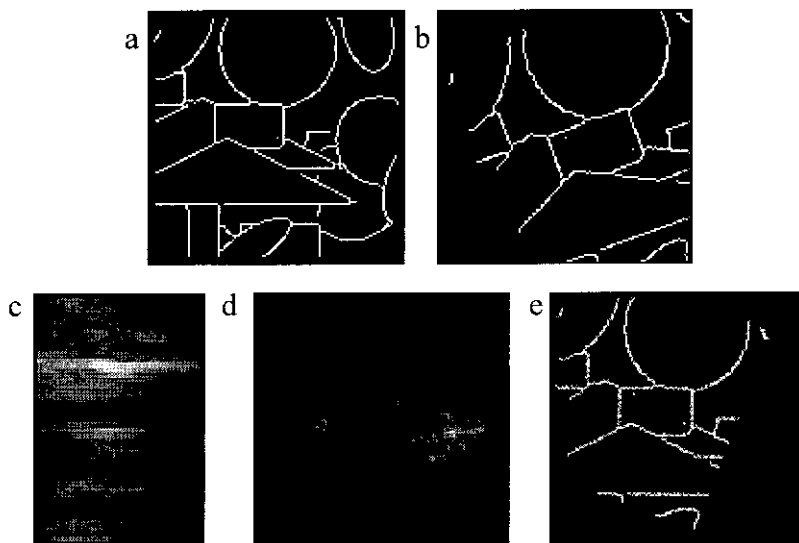
$$N_c \approx (C_{N,2})^2 = \left( \frac{N!}{2!(N-2)!} \right)^2 = \left( \frac{N \cdot (N-1)}{2} \right)^2 \approx N^4$$

Para casamientos entre segmentos aislados se obtiene:

$$N_c = N_1 \cdot N_2 \approx N^2$$

Por tanto, la simplificación de las estructuras empleadas para la determinación de los casamientos locales reduce de forma significativa la magnitud de los cálculos que deben ser realizados.

A modo de ejemplo, se muestran, en la figura 4.12, los resultados del registro de dos imágenes de bordes mediante la acumulación de transformaciones basadas en el casamiento de parejas de segmentos, sobre una matriz de acumulación 4D. El modelo de transformación contempla un escalado global, un giro y las dos traslaciones. Las imágenes 4.12c y 4.12d muestran, respectivamente, los planos de corte (que no proyecciones) de escalado-giro y traslaciones, correspondientes al máximo global de la matriz de acumulación 4D. Puesto que no se ha considerado el sesgo y los escalados se han supuesto iguales para el modelo de transformación, únicamente es necesario barrer una de las coordenadas (el escalado) del espacio de acumulación para poder determinar las transformaciones completas.



**Figura 4.12.** Registro de dos imágenes sintéticas mediante acumulación 4D en base a casamientos de parejas de segmentos, utilizando la superposición como función de peso. (a) Imagen sintética y (b) par de registro obtenido mediante una transformación afin de (a). Planos de corte en (c) escalado-ángulo y (d) traslaciones. (e) Resultado de la reconstrucción de la imagen 4.12b mediante la transformación correspondiente al máximo de acumulación detectado.

Las precisiones y rangos de los parámetros del modelo han sido los siguientes:

$$\begin{array}{ll}
 0.5 \leq s < 1.5 & ds = 0.03 \\
 -30^\circ \leq \alpha < 30^\circ & d\alpha = 1.2^\circ \\
 64 \text{ píxeles} \leq t_x, t_y < 64 \text{ píxeles} & dt_x = dt_y = 1.5 \text{ píxeles}
 \end{array}$$

La matriz de acumulación utilizaba doble precisión (8 bytes) para los valores de las celdillas, lo que resulta en una cantidad total de memoria de almacenamiento de:

$$\Omega = 34 \cdot 51 \cdot 86 \cdot 86 \cdot 8 = 102 \text{ Mbytes}$$

Requerimientos de memoria de estos ordenes son los que principalmente limitan la aplicabilidad de las técnicas de acumulación a la restauración de transformaciones con más de dos o tres parámetros. Aun así, los resultados son buenos, como puede verse en la imagen 4.12e, registrada mediante los parámetros del máximo detectado.

#### **4.6. Valoración y limitaciones del método de acumulación**

La acumulación es una técnica muy potente cuando se necesita analizar un conjunto de medidas grande para la determinación de una o varias funciones compatibles con el conjunto. Sin embargo, el crecimiento exponencial de los recursos necesarios, en función del número de parámetros, limita la complejidad del modelo de función que puede utilizarse.

Es posible emplear técnicas para soslayar esa dificultad, como la utilización de acumulación sobre proyecciones bidimensionales del espacio de acumulación (sección 4.3). De esta forma se reduce la cantidad de memoria necesaria, lo cual permite aumentar los límites y la precisión de cada uno de los parámetros de la función, pero con el inconveniente de generar ruido de acumulación, que puede dificultar la detección del máximo o generar acumulaciones espurias que oculten al máximo real.

La determinación de la precisión de los parámetros de la transformación para cada método de registro es también crítica para el correcto funcionamiento del método de acumulación y dificulta su aplicación práctica para la automatización del proceso de registro.

No obstante, cuando el número de parámetros que hay que determinar es pequeño y se está cerca del máximo, el método de acumulación ofrece buenos resultados. Es el caso del refinamiento de las transformaciones de registro para imágenes visible-SAR (sección 8.2). También se puede aplicar la técnica de acumulación, con funciones de peso basadas en la superposición de los segmentos casados, para el registro de imágenes de bandas espectrales cercanas.

Por todo ello, y aunque la acumulación es una herramienta poderosa para algunos casos de registro, su utilización para el registro de imágenes de bandas espectrales alejadas no resulta completamente satisfactoria, debido principalmente a la imposibilidad de encontrar una función de evaluación local de casamientos que posea, después de la acumulación, utilidad global para la determinación de la mejor transformación de registro. La utilización de medidas globales convierten en innecesarias las acumulaciones, y no son, por tanto, compatibles con el método descrito en este capítulo.

En definitiva, para el registro genérico de imágenes de bandas espectrales alejadas debe buscarse algún método alternativo, como el registro basado en la calidad de casamientos locales de segmentos descrito en el capítulo siguiente.

## **5. Registro automático basado en el casamiento de conjuntos de segmentos lineales**

En el capítulo 4 se describía el primero de los dos métodos desarrollados en este trabajo para la determinación de la transformación global de registro a partir del casamiento local de estructuras formadas por segmentos. El otro método, basado en la definición de un criterio de calidad de casamiento entre segmentos para evaluar de forma global cada una de las transformaciones candidatas, se explica en este capítulo.

### **5.1. Introducción**

La evaluación de una transformación geométrica candidata para el registro de dos imágenes de bordes precisa de un procedimiento capaz de medir la correspondencia entre la imagen de bordes fuente transformada y la imagen de bordes destino. La cantidad de información contenida en una imagen de bordes es menor, y presenta más dispersión, que la existente en la imagen de grises de la que proviene, por lo que métodos basados en correlación y similares no son válidos para esta tarea. Por otra parte, puesto que las transformaciones de registro candidatas han sido determinadas mediante casamientos locales entre estructuras formadas por segmentos detectados en las imágenes de bordes (capítulo 3), es de gran interés utilizar esos mismos conjuntos de segmentos como base para la evaluación de las transformaciones de registro candidatas.

Para medir la posible correspondencia entre los dos conjuntos de segmentos, será necesario, en última instancia, disponer de una función capaz de evaluar el grado de coincidencia existente entre uno de los segmentos del conjunto fuente, una vez que se le ha aplicado la transformación de registro candidata, con otro segmento del conjunto destino. Se han propuesto varias técnicas para evaluar esa coincidencia, sobre todo orientadas a visión estereoscópica, pero ninguna de ellas resulta adecuada para su aplicación al registro de imágenes de bandas espectrales alejadas. La razón principal es que prácticamente todas ellas aceptan implícitamente que los bordes obtenidos serán muy similares en las dos imágenes, condición que no se cumple cuando los bordes proceden de imágenes de diferentes bandas espectrales. En el apartado 5.2.1 se hace una breve descripción de los principales métodos propuestos hasta ahora.



Para su aplicación a los métodos automáticos de registro descritos en esta memoria, se han desarrollado dos procedimientos de valoración local del casamiento entre segmentos. El primero de ellos se basa en la determinación de la superposición entre los segmentos fuente y destino, y puede utilizarse para el registro de imágenes del mismo tipo o de bandas espectrales cercanas, tal y como se ha mostrado en el capítulo 4. El segundo, válido para cualquier tipo de imágenes, emplea un criterio de calidad de casamiento entre segmentos basado en la premisa de que los únicos atributos invariantes entre imágenes de bandas espectrales diferentes son las rectas definidas por los segmentos lineales detectados en ellas. Este último método tiene también la ventaja de que puede emplearse para detectar casamientos, además de para evaluarlos.

En ambos casos se supondrá que el segmento de la imagen fuente ha sido transformado sobre la imagen destino, donde se compara con su segmento destino correspondiente.

## **5.2. Calidad de casamiento de dos segmentos lineales**

La evaluación de la posible correspondencia entre dos conjuntos de segmentos, debe basarse, en última instancia, en una función capaz de medir el grado de coincidencia existente entre dos segmentos individuales de cada uno de los conjuntos. En esta sección se describen las técnicas existentes y los dos nuevos métodos que hemos desarrollado para ello.

### **5.2.1. Antecedentes**

En la literatura se han descrito diferentes métodos para la comparación y el casamiento de segmentos lineales rectos o, en general, de curvas planas. La mayoría de ellos, sin embargo, están orientados a la estereoscopia o al registro del mismo tipo de imágenes.

El método propuesto por Zhang y Gerbrands [29] está orientado al casamiento de curvas cerradas, aunque puede aplicarse también para la comparación de segmentos rectos. La técnica propuesta consiste en medir la distancia mínima de cada punto de la curva fuente a la curva destino. Se contabiliza el número de puntos que se encuentran a una distancia menor que un cierto umbral y se divide por el número total de puntos de la curva destino. Ese cociente refleja qué proporción de puntos de la curva ha sido correctamente puesto en correspondencia. La utilización de este método para el casamiento de segmentos, no

obstante, valoraría más aquellos casamientos en los que el segmento destino fuese de igual o menor tamaño que el fuente, con lo que se puntuaban mejor, en general, los escalados mayores que uno. Por otra parte, el incremento de longitud del segmento fuente o del destino (como resultado de posibles diferencias en el proceso de extracción o ya presentes en las propias imágenes fuente), tendrá efecto sobre el valor del casamiento, lo cual tampoco es deseable para el registro de imágenes diferentes. Finalmente, aunque los segmentos se encuentren sobre la misma recta, en cuanto sus extremos más cercanos estén alejados por encima del umbral fijado, la calidad de casamiento será cero, algo que no es en absoluto deseable (dado que suponemos que la coincidencia de las rectas definidas por los segmentos es lo único que puede utilizarse para el registro de conjuntos de segmentos procedentes de imágenes de diferentes bandas).

En el artículo de Horaud y Skordas [38] la evaluación de la correspondencia se realiza en función de las diferencias de contraste de cada segmento, las de longitud, las diferencias de orientación y las de número de líneas vecinas existentes. De todas ellas, solamente las de orientación serían utilizables para el registro de imágenes de bordes de bandas espectrales separadas. Medioni y Nevatia [33] también hacen uso del contraste de los bordes, que no se conserva en imágenes diferentes.

El método descrito por Taylor y Kriegman [37] se basa en la suma de las distancias desde cada uno de los puntos del segmento destino hasta la recta predicha por la transformación candidata. Sin embargo, no considera el ángulo que forman la recta y el segmento, con lo que un segmento completamente perpendicular a la recta y con su centro contenido en ella será valorado igual que un segmento paralelo a la recta situado a una distancia de  $l/4 + 1/2$ , donde  $l$  es la longitud del segmento fuente transformado. En realidad, la valoración del casamiento entre dos segmentos perpendiculares debe ser directamente cero, ya que no es posible que se correspondan con el mismo borde físico de la escena.

En el artículo de Sull y Ahuja [36] se utiliza de nuevo la distancia perpendicular a la recta destino, pero, en este caso, únicamente para los puntos extremos del segmento, que no se conservan en el proceso de extracción. Tampoco se tiene en cuenta ninguna consideración angular.

El procedimiento empleado por Behzad y Behrooz Kamgar-Parsi [40] supone que las longitudes de los segmentos a casar deben ser iguales, es decir, no considera la posibilidad de que existan diferencias entre segmentos correspondientes, causadas por el proceso de

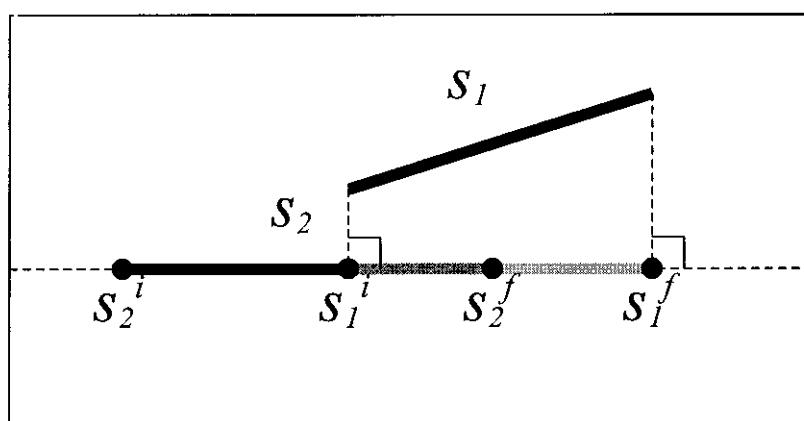
extracción o debidas a diferencias de los propios bordes en las imágenes iniciales. Por tanto, este método tampoco es aplicable al registro de imágenes de diferente naturaleza.

Finalmente, en el procedimiento desarrollado por Gros, Bournez y Boyer [35], aunque se tiene en cuenta la diferencia de orientación de los segmentos, vuelve a emplearse el cociente entre sus longitudes como parámetro para la correspondencia, lo que invalida también a su método para el registro de imágenes de diferentes bandas espectrales.

### 5.2.2. Calidad basada en superposición

La valoración de los casamientos en función de la longitud de superposición puede emplearse para el registro de imágenes de bandas próximas o que hayan sido obtenidas con el mismo dispositivo. La combinación de este criterio de valoración con los métodos de extracción de posibles transformaciones (descritos en los capítulos 3 y 4) constituye un procedimiento robusto para el registro de ese tipo de imágenes, siempre y cuando las deformaciones estén acotadas y el número de parámetros a restaurar sea pequeño.

Nuestro método de evaluación basado en superposición se inicia con la proyección del segmento fuente transformado  $s_1$  sobre la recta definida por el segmento destino  $s_2$ , como se muestra en la figura 5.1. Después se utilizan los parámetros de recta de cada uno de los vértices de los segmentos para obtener la longitud de superposición.



**Figura 5.1.** Longitud de superposición de dos segmentos. Para evaluarla se proyecta el segmento fuente transformado  $s_1$  sobre la recta definida por el segmento destino  $s_2$ .

La función utilizada para evaluar esa longitud de superposición en el espacio de la imagen destino es la siguiente:

$$S(s_1, s_2) = \frac{|t_2^f - t_2^i| + |t_1^f - t_1^i| - |t_2^f - t_1^f| - |t_2^i - t_1^i|}{2} \cdot \|s_2\| \quad (5.1)$$

donde los superíndices  $i$  o  $f$  indican si el extremo del segmento es el inicio o el final del mismo, lo cual se decide después de calcular los parámetros de recta,  $t_n^m$ , de cada punto  $s_n^m = (x_n^m, y_n^m)$  sobre la recta definida por  $s_2$ , que pueden obtenerse mediante:

$$t_n^m = \frac{x_n^m - x_0}{v_x} \quad \text{ó} \quad t_n^m = \frac{y_n^m - y_0}{v_y} \quad (5.2)$$

$x_0, y_0$  son las coordenadas de un punto arbitrario de la recta definida por el segmento  $s_2$  y  $v_x, v_y$  definen su vector director,  $\hat{s}_2$ :

$$\hat{s}_2 = \frac{\overline{s_2}}{\|s_2\|} = (v_x, v_y)$$

La expresión paramétrica de la recta es la siguiente:

$$\begin{cases} x = v_x t + x_0 \\ y = v_y t + y_0 \end{cases} \quad (5.3)$$

La función de superposición  $S(s_1, s_2)$  (5.1) es similar a la utilizada por Zhang [34], pero en aquella el resultado no indica si existe o no superposición, por lo que esta cuestión debe decidirse antes de aplicar la función de medida. El resultado de  $S$ , cuando no hay superposición, es directamente negativo.

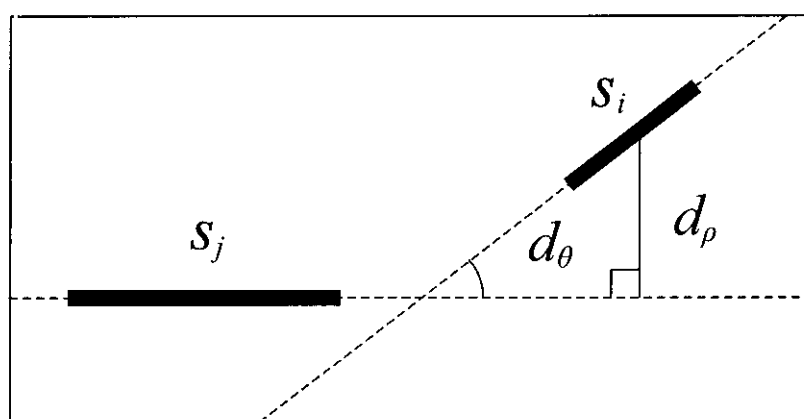
La función (5.1) puede emplearse directamente como función de peso para el registro de imágenes mediante procedimientos de acumulación (capítulo 4).

### 5.2.3. Calidad basada en distancia

Para evaluar la posible correspondencia entre un segmento fuente transformado  $s_i$  y un segmento destino  $s_j$  procedentes de imágenes de bandas espectrales alejadas, es necesario desarrollar un criterio que considere nuestra premisa básica para el registro de los bordes de imágenes diferentes, esto es, que tenga en cuenta que la única característica conservada entre segmentos correspondientes es la recta definida por cada segmento. Un modo de hacerlo consiste en definir una función capaz de evaluar la distancia del segmento fuente transformado a la recta definida por el segmento destino.

La principal ventaja que tiene la definición de una función de distancia de este tipo es que, además de servir para valorar un casamiento determinado, también puede emplearse para encontrar el segmento destino con el que más probablemente casará un segmento fuente transformado mediante una función de registro candidata. En este sentido, todos los criterios de valoración de casamientos basados en superposición o similares (como los descritos en 5.2.1 y 5.2.2) pueden emplearse para el diseño de procedimientos de evaluación, pero no para la búsqueda de posibles casamientos.

Para la definición de una función de distancia desde un segmento fuente transformado,  $s_i$ , hasta un segmento destino,  $s_j$ , basada en la comparación de las rectas que generan, hemos utilizado sus diferencias de orientación y la distancia perpendicular desde el centro de  $s_i$  hasta la recta definida por  $s_j$  (figura 5.2). El valor de la distancia de  $s_i$  a  $s_j$ , además, debe disminuir según se acerque  $s_i$  a la recta definida por  $s_j$ , y debe ser cero si está contenido en ella.



**Figura 5.2.** La distancia del segmento fuente transformado  $s_i$  al segmento destino  $s_j$  depende de su diferencia de orientación,  $d_\theta(s_i, s_j)$  y de la distancia desde el centro de  $s_i$  hasta la línea definida por  $s_j$ ,  $d_\rho(s_i, s_j)$ .

En base a esas condiciones, hemos definido la siguiente función de distancia:

$$D_{\rho\theta}(s_i, s_j) = \frac{1}{\sqrt{2}} \left( \left( \frac{d_\rho(s_i, s_j)}{d_{\rho\max}} \right)^2 + \left( \frac{d_\theta(s_i, s_j)}{d_{\theta\max}} \right)^2 \right)^{1/2} \quad (5.5)$$

donde  $d_\rho(s_i, s_j)$  es la distancia desde el centro de  $s_i$  hasta la línea definida por  $s_j$ ,  $d_\theta(s_i, s_j)$  es la distancia angular (diferencia de orientación entre  $s_i$  y  $s_j$ ), y  $d_{\rho\max}$  y  $d_{\theta\max}$  son los valores límite para  $d_\rho$  y  $d_\theta$  respectivamente, utilizados en la normalización de la función.

Como puede verse, el valor de  $D_{\rho\theta}$  es 0 si el segmento fuente transformado está completamente contenido dentro de la recta definida por el segmento destino, ya que en ese caso  $d_\rho$  y  $d_\theta$  se anulan. Si el valor de la distancia es mayor que 1 se supone que los segmentos no se corresponden, ya que en tal caso habrán superado los valores límite especificados,  $d_{\rho\max}$  y  $d_{\theta\max}$ .

El factor de normalización  $1/\sqrt{2}$  se incluye para que no pueda darse el caso en el que, aún dentro de las tolerancias fijadas, existan puntos  $(d_\rho, d_\theta)$  que se consideren fuera de límites. De esta forma, el punto  $(d_{\rho\max}, d_{\theta\max})$  está justo en el límite de lo que se considera casamiento. Como contrapartida, existirán puntos en los que una de las tolerancias pueda ser superada, cuando la otra sea muy pequeña (por ejemplo, para  $d_\rho = 0$ ,  $D_{\rho\theta}$  se hace 1 para  $d_\theta$  mayor que  $d_{\theta\max}$ ). El factor de normalización es principalmente una cuestión cosmética, ya que es fácil establecer una correspondencia entre los límites máximos con o sin constante de normalización y ajustar estas últimas en consecuencia. Los valores máximos de  $d_\rho$  y  $d_\theta$ , para la normalización elegida, serán entonces:

$$D_{\rho\theta}(s_i, s_j) = 1 \Rightarrow \begin{cases} d_\rho(s_i, s_j) = 0 \Leftrightarrow |d_\theta(s_i, s_j)| = \sqrt{2} \cdot d_{\theta\max} \\ d_\theta(s_i, s_j) = 0 \Leftrightarrow |d_\rho(s_i, s_j)| = \sqrt{2} \cdot d_{\rho\max} \end{cases} \quad (5.6)$$

Los valores utilizados usualmente en nuestras pruebas de registro (imágenes visible-IR de 320x240 píxeles) son los siguientes:

$$d_{\rho\max} = 5 \text{ píxeles}; \quad d_{\theta\max} = 0.2 \text{ rad} \quad (5.7)$$

En algunas ocasiones, en las que la precisión en el casamiento es crítica (como en la utilización del registro para el reconocimiento de formas, en la sección 8.3.1), pueden especificarse límites más severos.

En base a esta función de distancia entre segmentos sería posible definir una función de distancia entre conjuntos de segmentos (por ejemplo, la suma de las distancias mínimas de cada uno de los segmentos fuente transformados a los segmentos del conjunto destino). Este enfoque, no obstante, presenta varios problemas. El principal de ellos es que la función de distancia no está acotada, con lo que segmentos fuente aislados, o sin correspondencia real en la imagen destino, aumentarán enormemente el valor de la distancia entre los conjuntos de segmentos. Por lo tanto, la evaluación de la correspondencia entre dos conjuntos de segmentos variará mucho en cuanto se incluyan unos pocos segmentos espurios a la imagen fuente (que pueden ser debidos a diferentes factores influyentes en el proceso de extracción). Análogamente, la adición de segmentos falsos a la imagen destino disminuirá el valor de la distancia entre los dos conjuntos de segmentos, puesto que será más probable que aparezca un segmento (correspondiente o no) en la posición final de cada segmento fuente transformado.

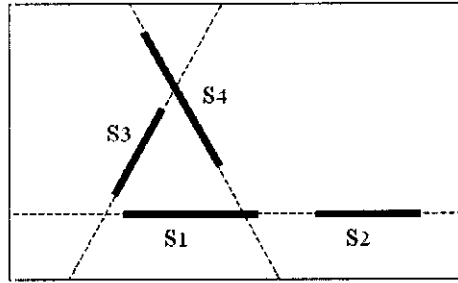
Por todo ello, resulta más conveniente definir una función de calidad de casamiento entre segmentos que, utilizando la distancia  $D_{\rho\theta}$  como base, asigne un valor mayor a las correspondencias más probables (con distancias próximas a cero), y se haga cero cuando la distancia supere un determinado valor. De esta manera, segmentos alejados o sin correspondencia real no afectarán a la calidad global del casamiento entre los conjuntos de segmentos, con lo que la función de evaluación será menos sensible a las diferencias que inevitablemente aparecen en la extracción de bordes y segmentos en imágenes de diferente naturaleza.

Definiremos por tanto, el factor de calidad de casamiento  $Q$  entre un segmento fuente transformado  $s_i$  y un segmento destino  $s_j$  como:

$$Q(s_i, s_j) = \begin{cases} 1 - D_{\rho\theta}(s_i, s_j) & \text{si } D_{\rho\theta}(s_i, s_j) \leq 1 \\ 0 & \text{si } D_{\rho\theta}(s_i, s_j) > 1 \end{cases} \quad (5.8)$$

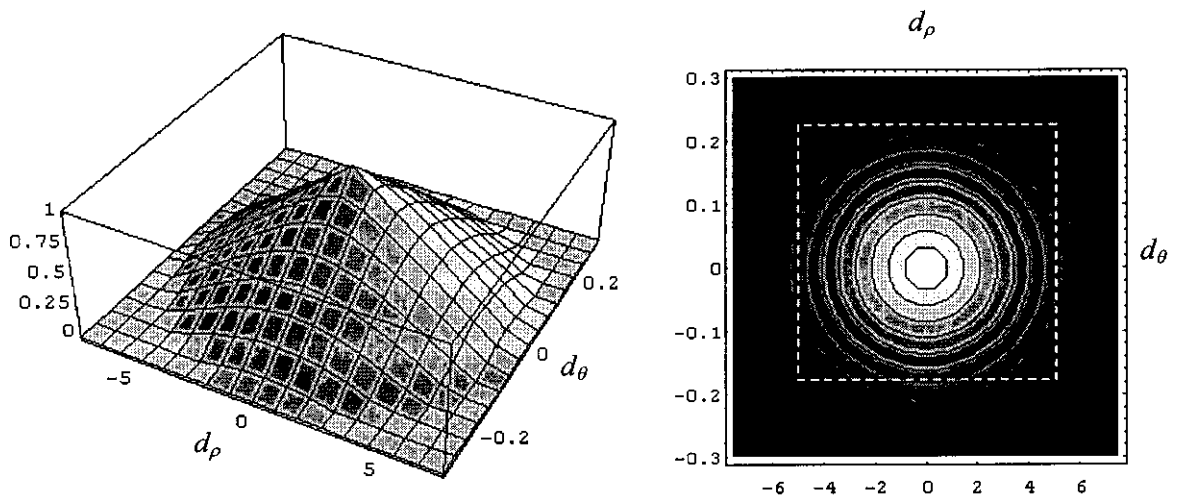
Esta función de calidad puede utilizarse también para la eliminación de segmentos redundantes (esto es, que dan lugar a rectas coincidentes) a la hora de calcular los posibles triángulos que forman, y que serán empleados para la determinación de las posibles transformaciones de registro. En efecto, cada dos segmentos colineales definen la misma

recta, por lo que generarán el mismo triángulo al agruparse con otros dos segmentos (no colineales), como se muestra en la figura 5.3. Utilizando el criterio de calidad, dos segmentos  $s_1$  y  $s_2$  se considerarán colineales si  $Q(s_1, s_2)$  y  $Q(s_2, s_1)$  tienen valores superiores a cero.



**Figura 5.3.** Redundancias en la generación de triángulos. Los segmentos  $s_1$  y  $s_2$  forman el mismo triángulo al agruparse con  $s_3$  y  $s_4$ . Los segmentos colineales próximos se corresponden frecuentemente con el mismo borde fragmentado de la escena real.

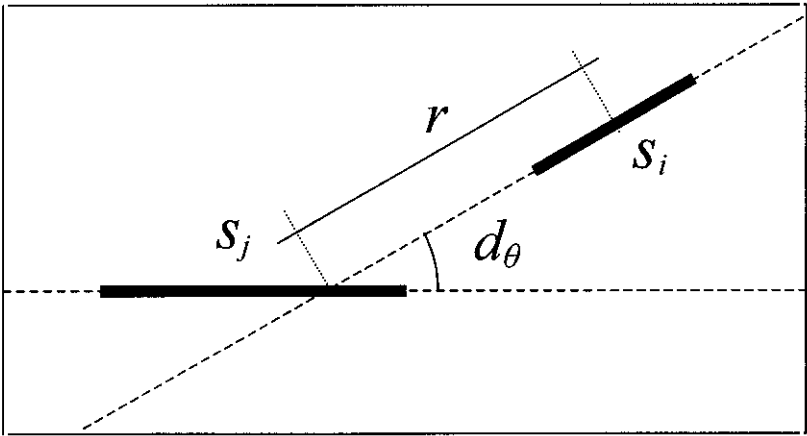
En la figura 5.4 puede observarse la forma de la función de calidad  $Q$  (5.8) en base a la variación de  $d_\rho$  y  $d_\theta$ , utilizando los valores de normalización para  $d_{\rho max}$  y  $d_{\theta max}$  indicados en (5.7), donde se aprecia el decremento lineal del valor del casamiento tanto en función de la distancia perpendicular como de la angular. La zona exterior al recuadro blanco en la imagen de curvas de nivel indica los puntos para los que se sobrepasa uno de los umbrales de normalización, pero no los dos, como se discutía en el párrafo sobre la normalización de  $D_{\rho\theta}$  (fórmula 5.5).



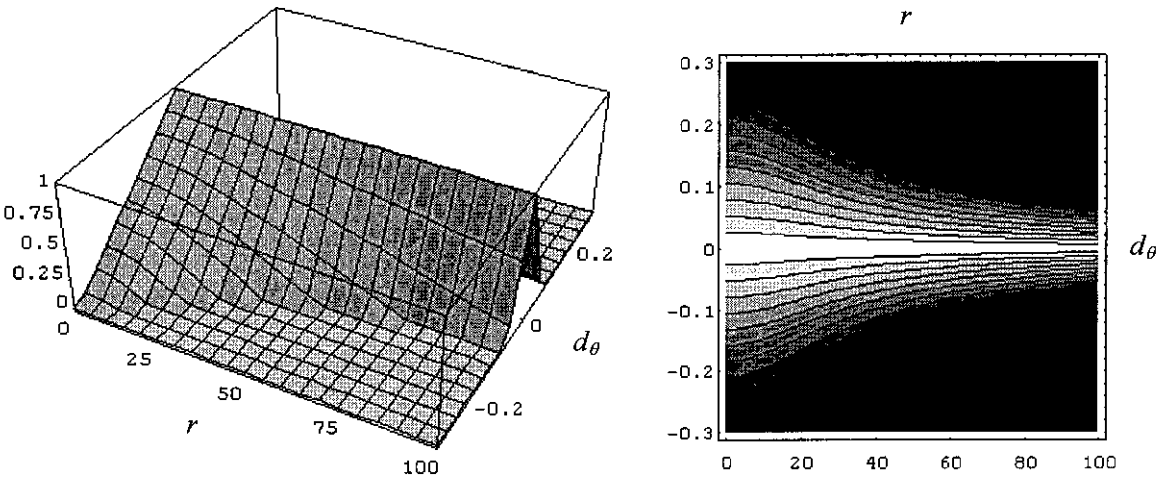
**Figura 5.4.** Variación de  $Q(s_i, s_j)$  en función de la diferencia de orientación,  $d_\theta(s_i, s_j)$  y de la distancia desde el centro de  $s_i$  hasta la línea definida por  $s_j$ ,  $d_\rho(s_i, s_j)$ . A la derecha se muestra la misma gráfica en forma de curvas de nivel, trazadas para valores múltiplos de 0.1.



Un caso interesante para analizar el comportamiento de la función de calidad  $Q(s_i, s_j)$  es aquél en el que  $s_i$  se desplaza a lo largo de la recta que lo contiene, como se muestra en la figura 5.5. El valor de  $Q$ , que dependerá de la distancia  $r$  del centro del segmento al punto de corte entre las rectas asociadas a los segmentos y del ángulo que formen, se muestra en la figura 5.6. Como puede observarse, para valores pequeños de  $r$  (es decir, cuando el centro del segmento  $s_i$  está muy cerca de la recta definida por  $s_j$ ) se admiten diferencias angulares grandes. Esa tolerancia a la diferencia de orientaciones disminuye según aumenta  $r$ , de modo que para segmentos  $s_i$  muy alejados, sólo existe casamiento si están contenidos en la recta definida por  $s_j$ .

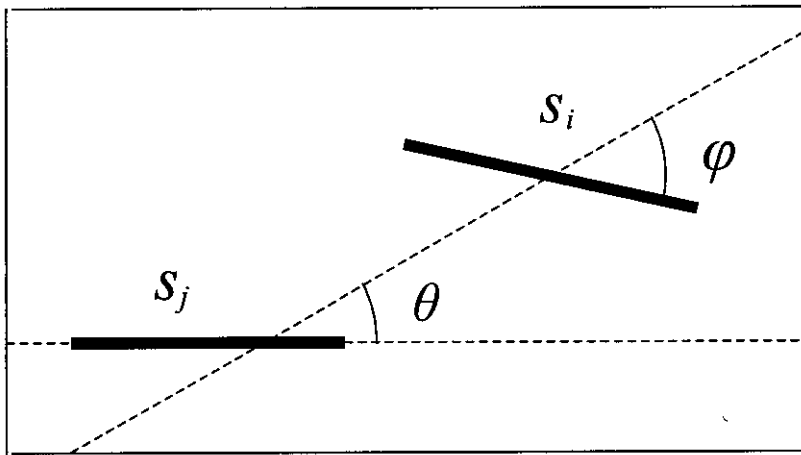


**Figura 5.5.** Esquema de la variación en la posición  $r$  del centro del segmento  $s_i$  a lo largo de la recta en la que esta contenido, la cual forma un ángulo  $d_\theta(s_i, s_j)$  con la recta que contiene al segmento destino  $s_j$ .

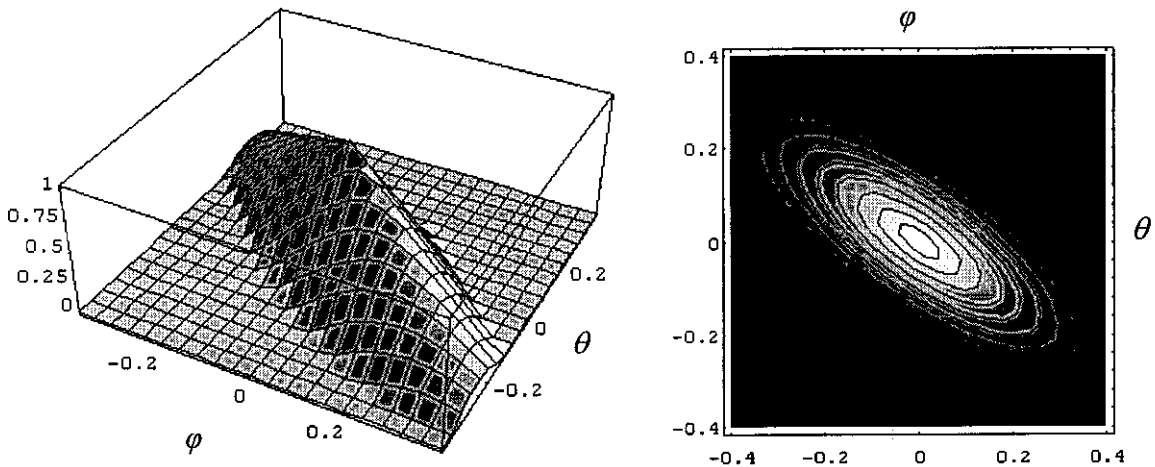


**Figura 5.6.** Valor de la calidad de casamiento entre los segmentos fuente y destino cuando la posición relativa de los mismos varía según el esquema de la figura 5.5. A la derecha, el mismo gráfico en forma de curvas de nivel, cada 0.1.

Otra situación de interés es aquélla en la que el centro del segmento fuente transformado se mantiene a distancia  $r$  fija del punto de corte entre las rectas y varía su orientación,  $\varphi$ , con respecto a la recta que contiene su centro, como se ilustra en la figura 5.7. La forma de  $Q$  en este caso (figura 5.8) indica que, para  $r$  fijo, la orientación del segmento fuente transformado,  $s_i$ , que maximiza la calidad es la que mantiene a  $s_i$  paralelo a la recta definida por  $s_j$ .



**Figura 5.7.** Esquema de la variación en la orientación  $\varphi$  del segmento  $s_i$  respecto a la orientación de la recta en la que está contenido su centro, la cual forma un ángulo  $\theta$  con la recta que contiene al segmento destino  $s_j$ . El centro del segmento  $s_i$  se mantiene a una distancia fija,  $r$ , del punto de corte de las rectas, como en el diagrama de la figura 5.5.



**Figura 5.8:** Valor de la calidad de casamiento entre los segmentos fuente y destino cuando la posición relativa de los mismos varía según el esquema de la figura 5.7, para una distancia  $r=20$  píxeles. A la derecha, el mismo gráfico en forma de curvas de nivel, cada 0.1.

### 5.3. Calidad de casamiento de dos conjuntos de segmentos lineales

A partir de la función para la evaluación de la calidad de casamiento de dos segmentos,  $Q(s_i, s_j)$  (fórmula 5.8), puede definirse una nueva función para medir la correspondencia entre dos conjuntos de segmentos,  $S_1$  y  $S_2$ . Esta función,  $Q(S_1, S_2)$ , permitirá evaluar transformaciones candidatas para el registro de las imágenes de las que han sido extraídos los conjuntos de segmentos. La transformación que genere el mayor valor de casamiento global será la que se seleccione como más probable para el registro de las imágenes.

Definimos la calidad de casamiento,  $Q(S_1, S_2)$ , entre el conjunto fuente transformado  $S_1$  y el conjunto destino  $S_2$  como la suma de las calidades de casamiento de cada uno de los segmentos fuente con el conjunto destino,  $Q(s_i, S_2)$ :

$$Q(S_1, S_2) = \sum_{s_i \in S_1} Q(s_i, S_2) \quad (5.9)$$

Al definir la calidad de casamiento entre conjuntos de esta forma, puede calcularse su valor independientemente de si se conocen o no las correspondencias existentes entre segmentos. Si, efectivamente, se dispone de la lista de asignaciones entre segmentos fuente y destino,  $L$ :

$$L = \{(s_i, s'_i), s_i \in S_1, s'_i \in S_2\} \quad (5.10)$$

puede escribirse:

$$Q(s_i, S_2) = \begin{cases} Q(s_i, s'_i) & \text{si } \exists (s_0, s'_0) \in L \mid s_0 = s_i \\ 0 & \text{en caso contrario} \end{cases} \quad (5.11)$$

y, por tanto:

$$Q_S(S_1, S_2) = \sum_L Q(s_i, s'_i) \quad (5.12)$$

Si, por el contrario, no se conoce  $L$ , definimos el segmento destino  $s'_i$  correspondiente a un cierto segmento fuente transformado,  $s_i$ , como aquel que maximiza el valor de la función de calidad  $Q$ :

$$Q(s_i, S_2) = \text{MAX}(\{Q(s_i, s_j), s_j \in S_2\}) \quad (5.13)$$

De esta forma, a la vez que se determina  $Q(S_1, S_2)$ , se obtiene la lista de casamientos  $L$ . En efecto:

$$L = \{(s_i, s'_i), s_i \in S_1, s'_i \in S_2 \mid 0 < Q(s_i, s'_i) = \text{MAX}\{\{Q(s_i, s_j), s_j \in S_2\}\}\}$$

Si el valor de  $Q(s_i, S_2)$ , tanto en (5.11) como en (5.13), es cero para un determinado segmento fuente transformado,  $s_i$ , significa que  $s_i$  no casa con ninguno de los segmentos del conjunto destino bajo la transformación aplicada.

Una vez determinada la lista de casamientos entre los conjuntos fuente y destino,  $L$ , puede emplearse cualquiera de las funciones de evaluación descritas en 5.2.1 o 5.2.2 para la obtención de un valor de registro global similar a  $Q(S_1, S_2)$  (5.9). Por ejemplo, el valor de superposición entre los dos conjuntos  $S_1$  y  $S_2$  en base a la función de superposición  $S(s_i, s_j)$  (5.1) sería:

$$Q_S(S_1, S_2) = \sum_L S(s_i, s'_i) \quad (5.14)$$

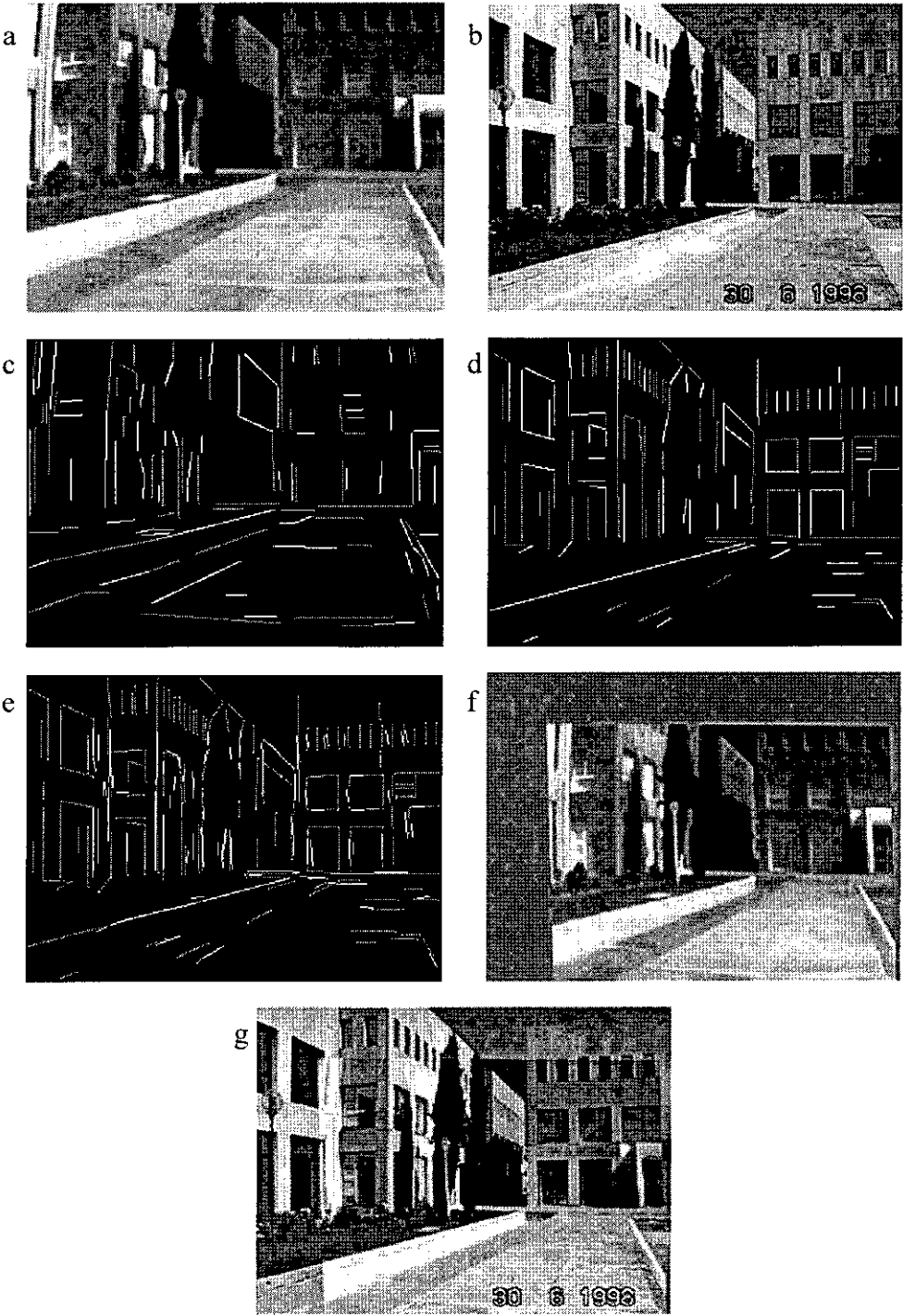
que puede utilizarse directamente para el registro de imágenes similares, como imágenes adquiridas mediante el mismo sensor o de bandas espectrales cercanas, pares estereoscópicos, etc.

En la figura 5.9 se muestra el resultado de la aplicación del método de registro basado en calidad de casamiento de conjuntos de segmentos a un par de imágenes IR-visible. Las imágenes 5.9c y 5.9d muestran los conjuntos de segmentos detectados en las imágenes. En blanco se indican los segmentos que han sido empleados para la generación de los conjuntos de triángulos (la utilización de los segmentos grises generaría redundancias en el conjunto de triángulos). En la imagen 5.9e se muestran los casamientos correspondientes a la transformación de registro óptima, que está definida por los siguientes parámetros:

$$\begin{aligned} s_x &= 0.826790 & s_k &= 0.038696 & t_x &= 21.763058 \text{ píxeles} \\ s_y &= 0.881568 & \alpha &= 0.008863 \text{ rad} & t_y &= 26.449606 \text{ píxeles} \end{aligned}$$

En la figura 5.9f puede verse el resultado de la aplicación de la transformación sobre la imagen IR. La fusión de la imagen IR registrada 5.9f con la visible 5.9b puede verse en la figura 5.9g. Como puede observarse en la imagen fundida, la transformación afín óptima no es capaz de registrar perfectamente las imágenes, ya que la deformación de desregistro incluye una componente de orden no lineal en la imagen IR. En efecto, líneas que deberían

aparecer rectas aparecen curvadas en la imagen 5.9a, debido a una distorsión en barrilete causada por la óptica de la cámara IR. La imagen fundida, en consecuencia, presenta algunos bordes dobles, principalmente en los extremos de la imagen IR.



**Figura 5.9.** Registro de una imagen IR con una visible de la misma escena. (a) Imagen IR, (b) imagen visible, (c) segmentos detectados en la imagen IR y (d) en la visible. (e) Segmentos de la imagen IR casados con la transformación afin óptima y (f) imagen IR registrada correspondiente. (g) Fusión de la imagen IR registrada con la imagen visible.

Los límites máximos permitidos para los parámetros de la transformación afín han sido:

$$0.7 < s_x, s_y < 1.3 \quad -0.1 \text{ rad} < \alpha < 0.1 \text{ rad}$$

$$-0.15 < s_k < 0.15 \quad -64.0 \text{ píxeles} < t_x, t_y < 64.0 \text{ píxeles}$$

El resto de los parámetros de configuración utilizados se muestra a continuación:

tamaño de las imágenes	320 x 240 píxeles
longitud mínima de los segmentos considerados	12 píxeles
anchura de la gaussiana para el extractor de bordes	5.0 píxeles
distancia máxima para la aproximación por segmentos	1.5 píxeles
ángulo mínimo permitido para triángulos	0.17 rad
área mínima permitida para triángulos	16 píxeles
$d_{pmax}$	5 píxeles
$d_{\theta max}$	0.2 rad

En el capítulo siguiente se describe el método que hemos desarrollado para poder obtener correcciones de registro de ordenes superiores al lineal (mediante el refinamiento de la transformación afín óptima calculada), de modo que puedan restaurarse también ese tipo de deformaciones de desregistro.



## **6. Cálculo de transformaciones de registro de ordenes superiores**

En los capítulos 4 y 5 se han descrito los métodos desarrollados en esta memoria para la determinación de la mejor transformación afín para el registro de dos imágenes de una misma escena. En este capítulo se explica el modo de refinar esa transformación para poder restaurar deformaciones de desregistro de ordenes superiores al lineal.

### **6.1. Introducción**

El registro de algunas imágenes puede requerir la utilización de modelos de deformación más complejos que las transformaciones afines resultantes de la aplicación de los métodos descritos en capítulos anteriores. Este tipo de modelos son necesarios cuando aparecen distorsiones no lineales debidas a la óptica o al método de funcionamiento de los sensores empleados para la adquisición de las imágenes.

En las ocasiones en las que se pueda calibrar previamente las imágenes en función de los parámetros de deformación de los dispositivos, podrán utilizarse transformaciones afines para el registro de las imágenes adquiridas por ellos. Sin embargo, en la mayoría de las ocasiones, únicamente se dispone del par de imágenes a registrar, sin ningún tipo de información adicional sobre el método de adquisición empleado. Por tanto, un método general de registro automático debe ser capaz de restaurar también las deformaciones de ordenes superiores al lineal que estén presentes en las imágenes.

Aunque en principio sería posible desarrollar un método que generase directamente transformaciones candidatas de ordenes superiores en función de casamientos de estructuras locales, es más práctico definir un método de refinamiento de la transformación de registro afín óptima en base a un modelo de deformación de orden superior. En la mayoría de los casos reales, además, las deformaciones no lineales son muy pequeñas frente a los escalados, rotaciones, traslaciones o sesgos, y no se pierde generalidad al considerarlas como refinamientos del modelo lineal.

### **6.2. Refinamiento de la transformación afín óptima**

Nuestro procedimiento para la obtención de una transformación de orden superior se basa en el refinamiento iterativo de la transformación afín de registro, utilizada como valor inicial de la transformación de orden superior. En la mayoría de los casos una corrección



de segundo o tercer orden es suficiente, y pueden usarse polinomios de esos grados como modelos de deformación. El método descrito a continuación utiliza un polinomio de segundo orden, aunque puede adaptarse de forma sencilla a cualquier otro tipo de funciones (por ejemplo, transformaciones que consideren efectos de perspectiva, para el registro de pares estereoscópicos).

El procedimiento de refinado emplea, en cada iteración  $k$ , una función de deformación  $F_k$  que transforma cada punto de la imagen fuente  $(x, y)$  en un punto de la imagen destino  $(x', y')$ . Suponiendo que  $F_k$  sea un polinomio de segundo orden, las ecuaciones de transformación son las siguientes:

$$\begin{aligned} x' &= a_0^k + a_1^k x + a_2^k y + a_3^k xy + a_4^k x^2 + a_5^k y^2 \\ y' &= b_0^k + b_1^k x + b_2^k y + b_3^k xy + b_4^k x^2 + b_5^k y^2 \end{aligned} \quad (6.1)$$

Para determinar los valores iniciales de los parámetros  $a_i^k$  y  $b_i^k$  deben compararse estas ecuaciones con las resultantes de aplicar el modelo de transformación afín empleado (fórmula 3.3):

$$\begin{aligned} x' &= m_0 x + m_1 y + m_2 \\ y' &= m_3 x + m_4 y + m_5 \end{aligned} \quad (6.2)$$

Se obtienen entonces los siguientes valores iniciales:

$$\begin{aligned} a_0^0 &= m_2 & a_1^0 &= m_0 & a_2^0 &= m_1 & a_3^0 &= a_4^0 &= a_5^0 &= 0 \\ b_0^0 &= m_5 & b_1^0 &= m_3 & b_2^0 &= m_4 & b_3^0 &= b_4^0 &= b_5^0 &= 0 \end{aligned} \quad (6.3)$$

En cada iteración,  $k$ , la transformación de registro  $F_k$  se aplica al conjunto de segmentos fuente. De esta forma, para cada segmento fuente,  $s = \{(x_1, y_1), (x_2, y_2)\}$ , se obtiene un nuevo segmento transformado,  $s' = \{(x'_1, y'_1), (x'_2, y'_2)\}$ . Si este segmento fuente casa con un segmento destino,  $t$ , las líneas que contienen  $s'$  y  $t$  deben ser muy próximas (suponiendo, por supuesto, que se emplea el criterio de casamiento basado en distancia definido en el capítulo 5). Por tanto, para calcular  $F_{k+1}$ , forzamos a  $s'$  a caer sobre la recta definida por  $t$ , proyectándolo sobre ella. Las coordenadas del segmento resultante,  $s'' = \{(x''_1, y''_1), (x''_2, y''_2)\}$ , se substituyen en (6.1) para obtener:

$$\begin{aligned}x_1'' &= a_0^{k+1} + a_1^{k+1}x_1 + a_2^{k+1}y_1 + a_3^{k+1}x_1y_1 + a_4^{k+1}x_1^2 + a_5^{k+1}y_1^2 \\x_2'' &= a_0^{k+1} + a_1^{k+1}x_2 + a_2^{k+1}y_2 + a_3^{k+1}x_2y_2 + a_4^{k+1}x_2^2 + a_5^{k+1}y_2^2\end{aligned}\tag{6.4}$$

y:

$$\begin{aligned}y_1'' &= b_0^{k+1} + b_1^{k+1}x_1 + b_2^{k+1}y_1 + b_3^{k+1}x_1y_1 + b_4^{k+1}x_1^2 + b_5^{k+1}y_1^2 \\y_2'' &= b_0^{k+1} + b_1^{k+1}x_2 + b_2^{k+1}y_2 + b_3^{k+1}x_2y_2 + b_4^{k+1}x_2^2 + b_5^{k+1}y_2^2\end{aligned}\tag{6.5}$$

Repitiendo el proceso para todos los segmentos casados, se generan dos conjuntos de ecuaciones (uno para los  $a_i^{k+1}$  y el otro para los  $b_i^{k+1}$ ), que pueden ser resueltos mediante descomposición en valores singulares [58] para obtener un nuevo conjunto de coeficientes,  $C_{k+1}$ . Este conjunto solución puede considerarse directamente como  $F_{k+1}$ . Sin embargo, para evitar cambios abruptos en la evolución de  $F$ , se mezclan estos nuevos coeficientes con los de la iteración anterior, empleando la expresión:

$$F_{k+1} = \chi \cdot F_k + (1 - \chi) \cdot C_{k+1}\tag{6.6}$$

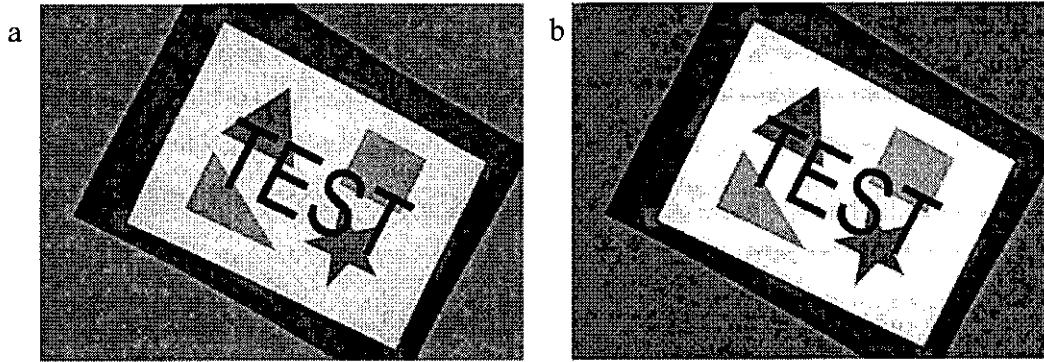
donde  $\chi$  es un número entre 0 y 1. Si las deformaciones a corregir son pequeñas puede utilizarse directamente el valor 0, ya que los casamientos entre segmentos serán los correctos, aunque su coincidencia espacial no sea perfecta. Para deformaciones mayores es posible que se hayan establecido casamientos iniciales incorrectos, y conviene dejar evolucionar el proceso de refinamiento de forma más lenta.

El valor de casamiento asociado a la nueva transformación  $F_{k+1}$  se calcula entonces por el método empleado para la determinación de la transformación afín óptima (superposición, calidad basada en distancia, etc). Si ese valor es mejor que el de la iteración anterior, se elige esta nueva transformación  $F_{k+1}$  como transformación de registro óptima, y se realiza una nueva iteración. El proceso se repite hasta que el incremento en el valor de casamiento es menor que un umbral prefijado (nosotros empleamos directamente 0). En cada iteración es posible emplear métodos de minimización (o maximización, dependiendo del criterio de valoración de casamientos que se esté utilizando), como el *simplex* [58], para acelerar la convergencia del proceso.

En la siguiente sección se muestran algunos ejemplos de la aplicación de esta técnica de refinamiento.

### 6.2.1. Ejemplos de aplicación del método de refinamiento

En la figura 6.1 se muestra el resultado de la aplicación, a la imagen de la figura 4.8a, de una corrección de tercer orden a la transformación afin de registro obtenida mediante acumulación (figuras 4.8h, i, j).



**Figura 6.1.** (a) Registro de la imagen 4.8a con la 4.8b mediante un polinomio de tercer orden, obtenido como refinamiento de la transformación afin determinada por el máximo de acumulación mostrado en las imágenes 4.8h, i, j. (b) Resultado de la fusión con la imagen 4.8b.

Los parámetros de la transformación afin eran:

$$\begin{aligned} s_x &= 0.773658 & s_k &= -0.002585 & t_x &= 15.211648 \text{ píxeles} \\ s_y &= 0.751265 & \alpha &= 0.512600 \text{ rad} & t_y &= -0.884711 \text{ píxeles} \end{aligned}$$

que se corresponden con los siguientes valores iniciales para los coeficientes del polinomio:

$$\begin{aligned} a_0^0 &= 15.211648 & b_0^0 &= -0.884711 \\ a_1^0 &= 0.67422175 & b_1^0 &= 0.37943661 \\ a_2^0 &= -0.37014636 & b_2^0 &= 0.65375412 \end{aligned}$$

Después del refinamiento, los coeficientes del polinomio de tercer orden resultantes son:

$$\begin{aligned} a_0 &= 15.07950447 & b_0 &= -1.57991829 \\ a_1 &= 0.67171751 & b_1 &= 0.38714079 \\ a_2 &= -0.38885533 & b_2 &= 0.65879779 \\ a_3 &= 0.00001602 & b_3 &= -0.00012891 \end{aligned}$$

$$a_4 = -0.00008188 \quad b_4 = -0.00001011$$

$$a_5 = 0.00003451 \quad b_5 = -0.00001340$$

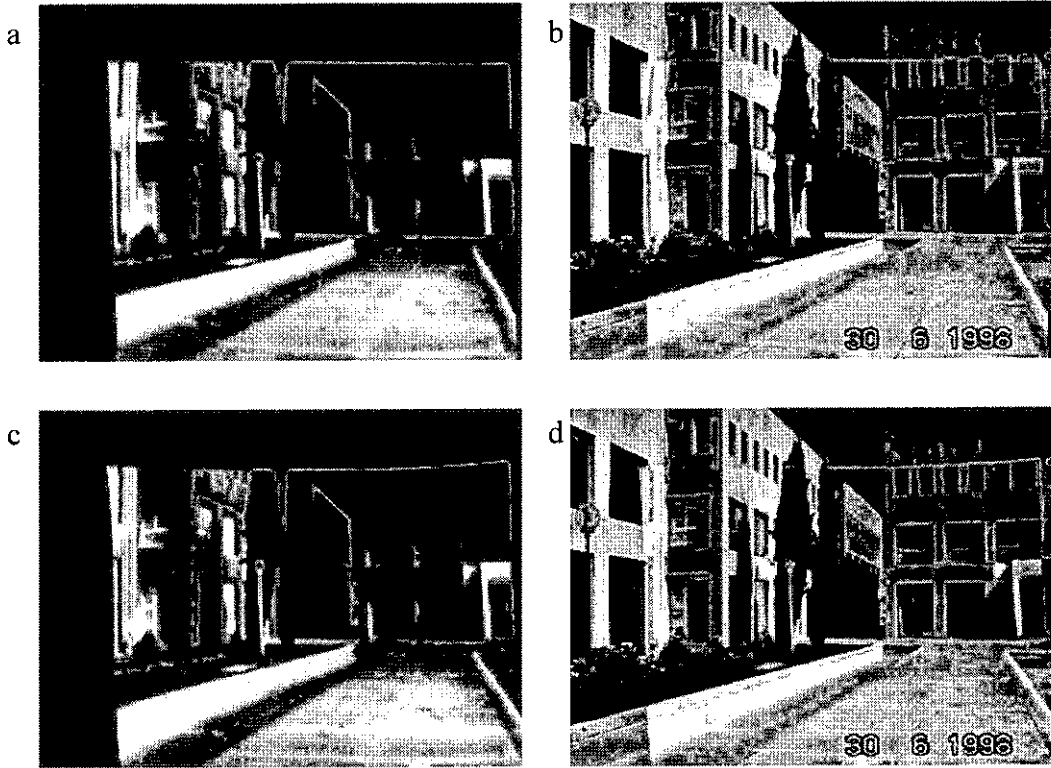
$$a_6 = 0.00000043 \quad b_6 = 0.00000104$$

$$a_7 = 0.00000000 \quad b_7 = 0.00000149$$

$$a_8 = -0.00000003 \quad b_8 = -0.00000039$$

$$a_9 = 0.00000077 \quad b_9 = 0.00000203$$

La corrección es pequeña, como se pone de manifiesto en la imagen 6.1a (muy parecida a la 4.11a), pero restaura completamente la deformación de distorsión de la cámara, como puede observarse en la imagen 6.1b, en la que no aparecen bordes duplicados (como ocurría en la 4.11, correspondiente al registro afín).



**Figura 6.2.** (a) Registro de la imagen IR 5.9a con una imagen visible de la misma escena (5.9b) mediante el método de casamiento de conjuntos de segmentos, y (b) fusión resultante. (c) Imagen IR registrada mediante la corrección de tercer orden de la transformación afín óptima y (d) fusión con la imagen visible.

En la figura 6.2 puede verse el resultado del refinamiento de tercer orden de la transformación de registro afín óptima calculada (mediante calidad de casamiento de segmentos) para el par de imágenes IR-visible de la figura 5.9. Comparando las imágenes 6.2a y 6.2c se aprecia claramente cómo la corrección de tercer orden corrige la distorsión

en barrilete de la imagen IR y suaviza las diferencias, debidas a la perspectiva, existentes en la parte inferior de las imágenes. Mediante la fusión se aprecia más claramente cómo la imagen registrada corregida en tercer orden (6.2d) se ajusta mejor a la imagen visible que la registrada mediante la transformación afín óptima (6.2b).

Los parámetros de la transformación afín eran:

$$\begin{array}{lll} s_x = 0.826790 & s_k = 0.038696 & t_x = 21.763058 \text{ píxeles} \\ s_y = 0.881568 & \alpha = 0.008863 \text{ rad} & t_y = 26.449606 \text{ píxeles} \end{array}$$

que se corresponden con los siguientes valores iniciales:

$$\begin{array}{ll} a_0^0 = 21.763058 & b_0^0 = 26.449606 \\ a_1^0 = 0.82675731 & b_1^0 = 0.00732770 \\ a_2^0 = 0.02629881 & b_2^0 = 0.88183528 \end{array}$$

Los coeficientes, después del refinamiento, se muestran a continuación:

$$\begin{array}{ll} a_0 = 21.35902359 & b_0 = 26.16350585 \\ a_1 = 0.82531492 & b_1 = 0.00623723 \\ a_2 = 0.02596700 & b_2 = 0.87005461 \\ a_3 = 0.00013294 & b_3 = -0.00008056 \\ a_4 = 0.00002353 & b_4 = -0.00001225 \\ a_5 = 0.00028811 & b_5 = 0.00010453 \\ a_6 = 0.00000034 & b_6 = 0.00000140 \\ a_7 = 0.00000149 & b_7 = -0.00000097 \\ a_8 = -0.00000002 & b_8 = -0.00000035 \\ a_9 = 0.000000230 & b_9 = 0.00000100 \end{array}$$

El método de refinamiento descrito en este capítulo, junto con el de determinación de transformaciones de registro a partir de casamientos de estructuras locales estudiado en el capítulo 5, constituyen un procedimiento genérico para el registro de cualquier tipo de imágenes. En el capítulo siguiente se desarrollará un método para evaluar cuantitativa y cualitativamente los resultados generados por este procedimiento de registro.

## 7. Evaluación global del registro

Una vez implementado nuestro método de registro para imágenes de bandas espectrales diferentes, es necesario evaluar sus resultados. Para ello hemos desarrollado dos técnicas principales: una basada en el proceso de descomposición de Gabor que empleamos para la fusión de imágenes, y la otra en el análisis del campo vectorial correspondiente a la función de deformación que registra las imágenes fuente y destino.

### 7.1. Evaluación basada en fusión de imágenes por descomposición de Gabor

Las diferencias en los atributos detectados en imágenes de diferente naturaleza dificultan la evaluación directa de los resultados de su registro. No obstante, puede utilizarse la misma información empleada por el registro para su propia evaluación. Esa información está contenida principalmente en los bordes de los objetos presentes en las imágenes. Como se ha visto anteriormente, no toda la información de bordes es directamente correlacionable entre imágenes de diferentes bandas, pero sí que es coincidente en gran parte de ellas, puesto que los bordes de las imágenes se corresponden en definitiva con bordes físicos reales de la escena. En consecuencia, si se suman o combinan dos imágenes incorrectamente registradas, debe producirse algún tipo de duplicación de bordes en la imagen resultante. El análisis de esta imagen combinada proporcionará, por tanto, información sobre la calidad de registro de las imágenes fuente y destino de las que procede.

La utilización de la fusión de imágenes [3-5] para la combinación y posterior evaluación de las imágenes registradas es preferible a otras técnicas más simples (sumas, promedios, etc), ya que tiende a preservar toda la información contenida en las imágenes de origen. Por otra parte, la fusión es menos sensible que otras técnicas más simples a la combinación de imágenes que presenten inversión de contraste, muy frecuente cuando se trabaja con diferentes bandas espectrales.

La fusión de las imágenes registradas permite además realizar una evaluación visual de la calidad del registro. El procedimiento de fusión que hemos elegido, desarrollado por Santamaría y Gómez [3], es muy sensible a la falta de registro, y en las zonas de las imágenes en las que éste no sea correcto aparecerán artefactos visuales (bordes dobles). En

la figura 7.1 se pone de manifiesto el efecto que tiene la falta de registro de dos imágenes sobre la imagen fundida resultante. La imagen 7.1a presenta una clara duplicación de bordes, debida a un pequeño desregistro entre las imágenes de partida. Cuando las imágenes están correctamente registradas (7.1b), los bordes dobles no aparecen.

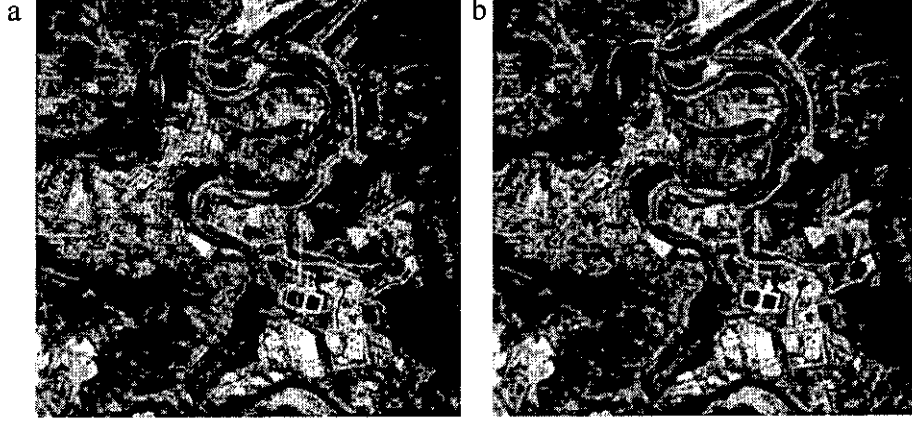
Sin embargo, aunque la inspección visual de la degradación que aparece en la imagen fundida permite realizar una valoración subjetiva del registro, debe cuantificarse de algún modo esa degradación para obtener una medida objetiva de la calidad de registro.

Se han desarrollado varios métodos para evaluar la calidad de imagen [60-63]. Generalmente se basan en el análisis del espectro de frecuencias espaciales modulado por la función de transferencia del ojo, con el fin de obtener una medida de la calidad visual de la imagen.

Sin embargo, la aplicación de ese tipo de técnicas para la evaluación del registro presenta algunas limitaciones. En efecto, una pequeña descorrelación global en los bordes de las imágenes registradas causada por traslación se comporta en el dominio de frecuencias como el producto con una senoide, que disminuye la energía de las componentes de algunas frecuencias del espectro. La curva de incremento de energía asociada a una degradación creciente de desregistro no es, por tanto, estrictamente creciente, y no puede utilizarse como medida proporcional al nivel de desregistro. Por otra parte, puesto que no se busca medir de forma absoluta la calidad de la imagen, sino únicamente el incremento de energía asociado a la distorsión geométrica de una de las imágenes en las proximidades de la deformación geométrica óptima, cálculos como el filtrado de ruido de las imágenes o la convolución con la función de transferencia del ojo, utilizados en ese tipo de criterios de calidad visual, no resultan necesarios.

Por todo ello, hemos desarrollado un nuevo método para la evaluación del registro, basado en el análisis de las frecuencias espaciales más altas exclusivamente.

La duplicación de bordes como consecuencia de un registro impreciso se corresponde, en definitiva, con la aparición de nuevas componentes de alta frecuencia espacial en la imagen fundida. En consecuencia, el peso relativo (o “energía”) de las altas frecuencias de una imagen fundida procedente de dos imágenes mal registradas debe ser, en general, mayor que la resultante de la fusión de las mismas imágenes correctamente registradas. Ese incremento, en el peso o valor relativo de las altas frecuencias espaciales, puede utilizarse entonces como medida del desregistro de las imágenes.



**Figura 7.1.** (a) Fusión de dos imágenes de bandas espectrales próximas ligeramente desregistradas, en la que se aprecia claramente la aparición de bordes dobles, debida a la falta de correspondencia espacial precisa entre las dos imágenes. (b) Fusión de las imágenes una vez que han sido correctamente registradas.

Por ello, y puesto que el procedimiento de fusión elegido [3] se basa en la descomposición en ondículas de Gabor [7, 13, 59], hemos utilizado la componente de Gabor de mayor frecuencia para realizar los cálculos del incremento de energía asociado al desregistro. De esta forma, además, la valoración de la calidad de registro se corresponde de forma más directa con la duplicación de bordes observable en la imagen fundida, puesto que se utiliza exactamente la misma zona del espectro que en aquella para el aislamiento de las altas frecuencias.

A continuación se describe brevemente el procedimiento de fusión utilizado, sobre el que se fundamenta nuestro método de evaluación del registro.

#### 7.1.1. Fusión de imágenes basada en ondículas de Gabor

Las funciones de Gabor bidimensionales tienen la forma siguiente [7]:

$$g_{x_0, y_0, f_n, \theta_m} = g_{0,0, f_n, \theta_m}(x, y) * \delta(x - x_0, y - y_0) \quad (7.1)$$

donde

$$g_{0,0, f_n, \theta_m}(x, y) = \exp\left(-\pi a^2 \left((x \cos \theta_m + y \sin \theta_m)^2 + (x \sin \theta_m - y \cos \theta_m)^2\right)\right) \cdot \exp(i(2\pi f_n (x \cos \theta_m + y \sin \theta_m))) \quad (7.2)$$



La función de Gabor está localizada tanto en el dominio espacial  $(x_0, y_0)$  como en el de frecuencias  $(f_n, \theta_m)$ , aquí en notación polar. El valor de  $a$  se corresponde con la anchura de banda radial de la función de Gabor, y es proporcional a la frecuencia radial  $f_n$  a la que se sintonice la función de Gabor:

$$a = \frac{1}{3} \sqrt{\frac{\pi}{\ln 2}} \cdot f_n \quad (7.3)$$

En el dominio de Fourier la función de Gabor (7.2) tiene la forma:

$$G_{0,0,f_n,\theta_m}(u,v) = \frac{1}{a^2} \exp\left(-\frac{\pi}{a^2} \left((u - f_n \cos \theta_m)^2 + (v - f_n \sin \theta_m)^2\right)\right) \quad (7.4)$$

La fusión de dos imágenes,  $I_1$  e  $I_2$ , basada en ondículas de Gabor se realiza en tres pasos principales: descomposición de Gabor de  $I_1$  e  $I_2$ , combinación de las componentes de Gabor resultantes y reconstrucción de la imagen final a partir de dichas componentes.

Para el proceso de descomposición seguimos el método empleado por Navarro, Nestares, Portilla y Taberner [7, 59], en el que se utilizan  $N = 4$  bandas de frecuencia y  $M = 4$  orientaciones:

$$f_n = \frac{2^{n-N}}{4} \text{ ciclos / píxel} \quad n = 1, \dots, N \quad (7.5)$$

$$\theta_m = \frac{\pi}{4}(m-1) \text{ radianes} \quad m = 1, \dots, M \quad (7.6)$$

Para agilizar la descomposición, ésta se realiza en el dominio espacial [59], mediante la utilización de un filtro resultante de la combinación de las ondículas pares correspondientes a las cuatro orientaciones posibles. En lugar de emplear máscaras diferentes para sintonizar diferentes frecuencias, y dada la utilización de frecuencias en intervalo de octavas, se realizan submuestreos sucesivos a tamaño mitad de la imagen a descomponer y se aplica el mismo filtro integrado. En la tabla 7.1 se muestran los coeficientes de la máscara empleada,  $M_{7 \times 7}$ .

Para la obtención de cada una de las componentes de Gabor,  $G_n$ , de una imagen  $I$ , únicamente es necesario, por tanto, convolucionar la imagen submuestreada al tamaño correspondiente a la frecuencia deseada,  $I_n$ , con la máscara integrada:

$$G_n = I_n * M_{7 \times 7} \quad (7.7)$$

A modo de ejemplo, se muestra en la figura 7.2 el resultado de la aplicación de la máscara  $M_{7 \times 7}$  a una imagen de 256x256 píxeles. La imagen 7.2b muestra la componente de Gabor de alta frecuencia,  $G_I$ , y la 7.2c su módulo. Como puede verse, el filtro únicamente responde a las zonas de la imagen en las que existen bordes asociados a frecuencias cubiertas por su ancho de banda.



**Figura 7.2.** Componente de Gabor de alta frecuencia,  $G_I$ , obtenida mediante el filtro  $M_{7 \times 7}$ . (a) Imagen original, (b) componente  $G_I$  (con signo) y (c) módulo de la componente de Gabor.

El recubrimiento espectral resultante del empleo de la máscara espacial integrada  $M_{7 \times 7}$  se aproxima bien al obtenido por la combinación de las ondículas pares en el dominio de frecuencias para las cuatro orientaciones elegidas, como se muestra en la figura 7.3. Emplearemos, por tanto, esta máscara integrada para el cálculo de las componentes de Gabor de alta frecuencia, sobre las que se basa el método de evaluación del registro desarrollado, descrito en la sección siguiente.

9.294605	4.966805	-7.966805	-5.589211	-7.966805	4.966805	9.294605
4.966805	-10.6888	-18.3112	-15.93361	-18.3112	-10.6888	4.966805
-7.966805	-18.3112	3.983402	37.93361	3.983402	-18.3112	-7.966805
-5.589211	-15.93361	37.93361	94.48962	37.93361	-15.93361	-5.589211
-7.966805	-18.3112	3.983402	37.93361	3.983402	-18.3112	-7.966805
4.966805	-10.6888	-18.3112	-15.93361	-18.3112	-10.6888	4.966805
9.294605	4.966805	-7.966805	-5.589211	-7.966805	4.966805	9.294605

Tabla 7.1. Máscara integrada,  $M_{7 \times 7}$ , utilizada para el filtrado de Gabor.

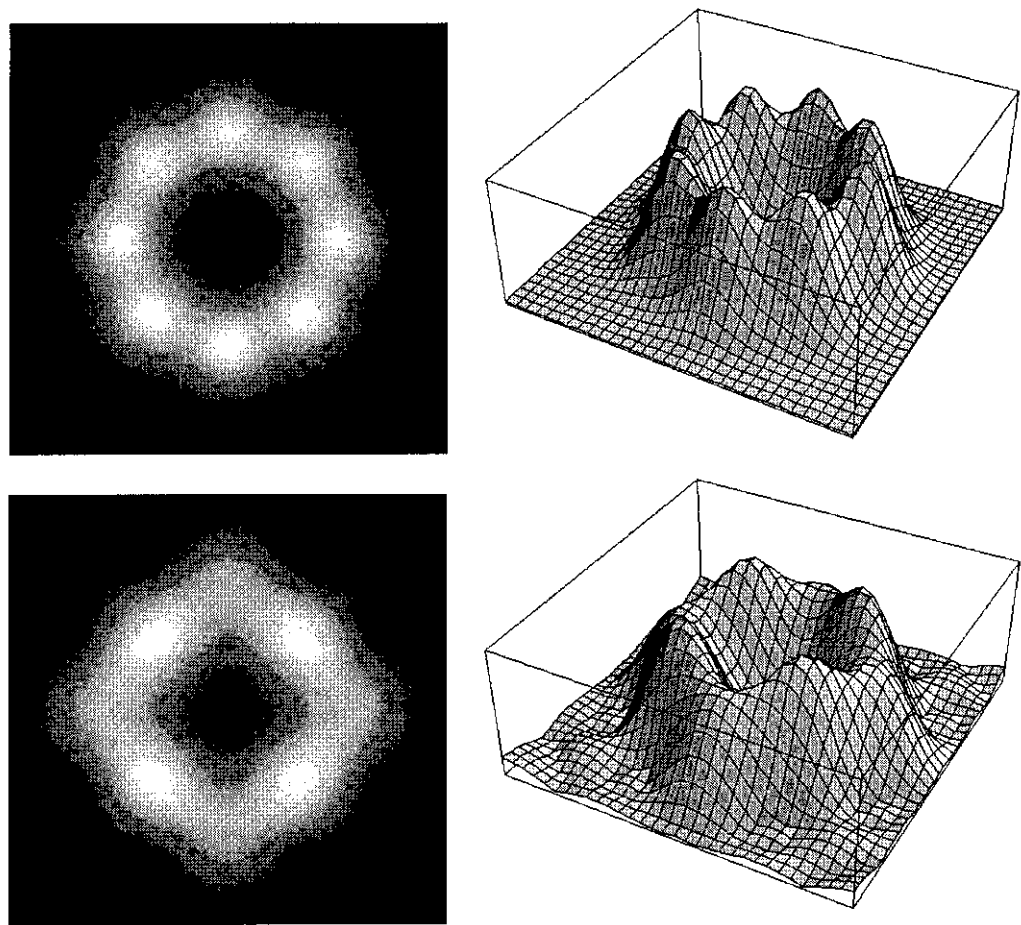


Figura 7.3. Recubrimiento del espectro de frecuencias para la componente de Gabor de frecuencia más alta. En la fila superior se muestra el recubrimiento combinado de las funciones de Gabor en frecuencia para las cuatro orientaciones necesarias. En la fila de abajo puede verse el recubrimiento correspondiente a la máscara integrada  $M_{7 \times 7}$ .

### 7.1.2. Criterio de evaluación del registro

Nuestro criterio de valoración de una transformación  $T$  para el registro de una imagen fuente  $I_1$  con una imagen destino  $I_2$ , se basa en el cálculo del incremento de energía producido en la banda de altas frecuencias de la imagen fundida resultante, comparado con la energía mínima asociada a la transformación de registro óptima  $T_{opt}$  para  $I_1$  e  $I_2$ :

$$\Delta E(T) = \frac{E(I_1, I_2, T)}{E_{min}} = \frac{E(I_1, I_2, T)}{E(I_1, I_2, T_{opt})} \quad (7.8)$$

La energía de altas frecuencias,  $E$ , asociada a una transformación  $T$  para el registro de dos imágenes,  $I_1$  e  $I_2$ , se determina mediante la integración, en el dominio de frecuencias espaciales, del producto del espectro de la fusión de las imágenes  $T(I_1)$  e  $I_2$  por el de la función de Gabor correspondiente a las frecuencias más altas. Por el teorema de Parseval sabemos que el resultado es equivalente al de la integración del producto de convolución, en el dominio espacial, de la imagen fundida con la máscara utilizada:

$$E(I_1, I_2, T) = \iint |\Phi(T(I_1), I_2) * M_{7x7}|^2 dx dy \quad (7.9)$$

donde  $T(I_1)$  es la imagen  $I_1$  transformada mediante la transformación de registro  $T$ , y  $\Phi(I, J)$  es el resultado de la fusión de las imágenes  $I, J$ .

La fusión de dos componentes de Gabor consiste en la extracción de la componente de máximo valor absoluto en cada píxel:

$$\Phi(G_1, G_2)(x, y) = \begin{cases} G_1(x, y) & \text{si } |G_1(x, y)| \geq |G_2(x, y)| \\ G_2(x, y) & \text{si } |G_1(x, y)| < |G_2(x, y)| \end{cases} \quad (7.10)$$

Puesto que se utilizan los mismos filtros para la realización de la fusión que para la extracción de la componente de Gabor de alta frecuencia, la extracción de las altas frecuencias de la imagen fundida es equivalente a la fusión de las componentes de alta frecuencia de las imágenes. Por lo tanto, la energía de más alta frecuencia puede calcularse de la siguiente manera:

$$E(I_1, I_2, T) = \iint |\Phi(T(I_1) * M_{7x7}, I_2 * M_{7x7})|^2 dx dy \quad (7.11)$$

con lo cual ya no resulta necesario fundir previamente las imágenes  $T(I_1)$  e  $I_2$ .

Para el cálculo de  $\Delta E$  mediante (7.8) es necesario disponer de la transformación de registro óptima,  $T_{opt}$ , que normalmente no se conoce. Únicamente cuando se apliquen deformaciones conocidas a imágenes (sintéticas o reales) previamente registradas, o cuando se disponga de una descripción completa de los parámetros que configuran los dispositivos de adquisición de las imágenes, se conocerá efectivamente la transformación de registro óptima. Por ello, en los casos prácticos reales en los que sea necesario evaluar la calidad del registro se realizará además un registro manual, y la transformación de registro obtenida,  $T_m$ , se considerará próxima a la transformación de registro óptima, y será utilizada como base para el cálculo de la energía mínima.

Para dar una significación más apropiada al método de evaluación empleado, es conveniente relacionar el incremento de energía asociado al desregistro con algún parámetro de tipo geométrico. Para ello, hemos desarrollado un criterio basado en el cálculo de los incrementos de energía correspondientes a deformaciones geométricas conocidas de la imagen registrada óptima,  $I_{1opt} = T_{opt}(I_1)$ , con los que será comparado el incremento de energía asociado a la transformación de registro que se desea evaluar.

Las deformaciones consideradas son, por simplicidad, únicamente traslaciones de la imagen  $I_{1opt}$ . Utilizándolas se crea un mapa de incrementos de energía, que, cerca del punto de traslación nula (correspondiente a la transformación óptima), es aproximadamente un mínimo con comportamiento lineal. Este mapa de incrementos depende directamente de las características de las imágenes consideradas, por lo que el comportamiento en los puntos alejados del punto de energía mínima es fuertemente no lineal y difícilmente modelable de forma genérica. Por ello, esta medida del desregistro sólo debe utilizarse cuando la calidad de la transformación de registro que se desea evaluar esté próxima a la de la transformación óptima. El mapa, en función del desplazamiento  $(\Delta t_x, \Delta t_y)$  se obtiene, a partir de (7.8), de la siguiente forma:

$$M(\Delta t_x, \Delta t_y) = \Delta E(T_{\Delta t_x, \Delta t_y}) \quad (7.12)$$

donde:

$$T_{\Delta t_x, \Delta t_y} = \begin{pmatrix} 1 & 0 & \Delta t_x \\ 0 & 1 & \Delta t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot T_m \quad (7.13)$$

Durante la creación del mapa es posible que el punto de energía mínima no coincida exactamente con la posición de traslación nula, es decir, que la transformación  $T_m$  elegida como transformación óptima puede ser mejorada con una pequeña traslación. En efecto, el registro manual no es necesariamente óptimo, y, en algunas ocasiones, el mínimo del mapa no se corresponderá exactamente con la traslación nula. La transformación asociada al punto de incremento mínimo se utilizará, por tanto, como transformación óptima:

$$T_{opt} \equiv T_{\Delta t_{x0}, \Delta t_{y0}} \mid M(\Delta t_{x0}, \Delta t_{y0}) = \text{MIN}\{M(\Delta t_x, \Delta t_y)\} \quad (7.14)$$

$$E_{min} = M(\Delta t_{x0}, \Delta t_{y0})$$

Puesto que la forma del mapa en los alrededores del mínimo raras veces será simétrica, existirán puntos a igual distancia del mínimo con incrementos de energía diferentes. Para obtener una medida más homogénea hemos definido, a partir del mapa de incrementos de energía, una función radial con el valor máximo encontrado a cada distancia:

$$\overline{\Delta E}(r) = \frac{\text{MAX}\{M(\Delta t_x, \Delta t_y) \mid \sqrt{\Delta t_x^2 + \Delta t_y^2} = r\}}{E_{min}} \quad (7.15)$$

De esta manera puede asociarse al incremento de energía de una transformación  $T$  dada, un valor de desplazamiento medio equivalente (en píxeles),  $\bar{r}_T$ :

$$\boxed{\bar{r}_T \equiv r \mid \overline{\Delta E}(r) = \overline{\Delta E}(T)} \quad (7.16)$$

Y cuanto mayor sea esta distancia equivalente, peor será la transformación  $T$  para el registro de las imágenes de entrada,  $I_1$  e  $I_2$ . Valores de  $\bar{r}_T$  próximos a cero indican que la transformación es adecuada para el registro de las imágenes.

En resumen, la calidad de una transformación  $T$  para registrar dos imágenes,  $I_1$  e  $I_2$ , se calcula de la siguiente manera:

- Se determina la transformación de registro óptima,  $T_{opt}$ . Sólo si la función de deformación entre  $I_1$  e  $I_2$  es conocida se sabe realmente el valor de  $T_{opt}$ . En cualquier otro caso se efectúa un registro manual para obtener una transformación de registro  $T_m$ , que se supone aproximada a  $T_{opt}$ .

- Se extrae el mapa de incrementos de energía para las imágenes  $T_m(I_1)$  e  $I_2$ . La transformación asociada al punto de mínima energía del mapa se considera como  $T_{opt}$ .
- A partir del mapa se extrae la curva de incrementos de energía en función de la distancia,  $\overline{\Delta E}(r)$ .
- Se calcula la transformación de registro,  $T$ , mediante el método automático que se desee evaluar.
- El incremento de energía  $\Delta E(T)$  de la fusión de las componentes de alta frecuencia de Gabor de  $T(I_1)$  e  $I_2$  se compara con la curva en función de la distancia,  $\overline{\Delta E}(r)$ , que determina el desplazamiento en píxeles  $\bar{r}_T$ , equivalente al desregistro resultante de aplicar  $T$  en lugar de  $T_{opt}$ .

### 7.1.3. Imágenes de prueba

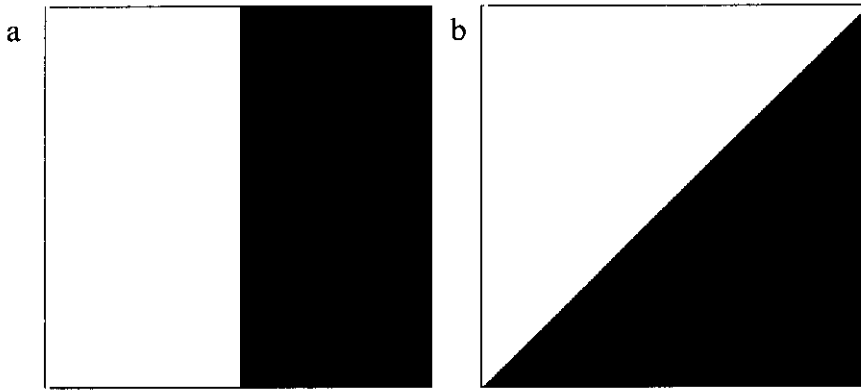
Para comprobar el funcionamiento del criterio de calidad de registro, se han utilizado varias imágenes de prueba: un borde ideal, una pareja de imágenes sintéticas, y un par de imágenes Daedalus de bandas espectrales diferentes, ya registradas.

#### 7.1.3.1. Borde Ideal

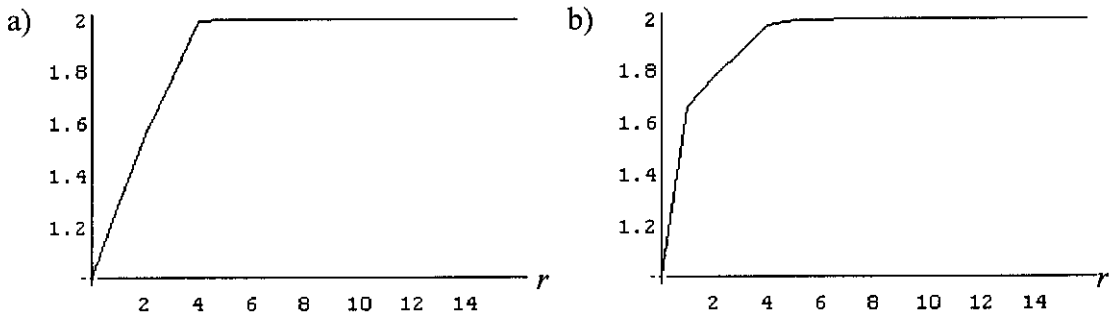
Las imágenes del borde ideal se han empleado para determinar el rango de aplicación del criterio de evaluación desarrollado. Puesto que la frecuencia a la que ha sido sintonizada la función de Gabor es de 1/4 ciclos/píxel, el máximo desplazamiento equivalente que puede medirse debe estar próximo al periodo asociado a esa frecuencia, es decir, unos cuatro píxeles. La anchura de banda,  $a$  (fórmula 7.3), extiende ligeramente el rango de frecuencias que intervienen en la determinación de la distancia equivalente, y la máscara  $M_{7 \times 7}$  utilizada como aproximación al filtro de Gabor sintonizado a 1/4 ciclos/píxel también tiene en cuenta otras frecuencias accesorias. Es importante, por tanto, comprobar experimentalmente que la máscara  $M_{7 \times 7}$  utilizada para extraer las componentes de Gabor genere los resultados esperados. Para ello se han empleado dos imágenes sintéticas con bordes ideales de gradientes orientados a 0 y 45° (mostradas en la figura 7.4), ya que esas

orientaciones determinan los ejes de simetría de la máscara  $M_{7 \times 7}$  en el dominio de Fourier (figura 7.2).

Las curvas de incremento de energía,  $\overline{\Delta E}(r)$ , obtenidas a partir del mapa de incrementos de cada imagen consigo misma, se muestran en la figura 7.5. En ellas puede verse que, efectivamente, el límite de percepción del filtro de Gabor integrado está alrededor de los cuatro píxeles de desregistro en traslación. Si hubiese que evaluar deformaciones mayores sería necesario, por tanto, utilizar filtros sintonizados a frecuencias menores.



**Figura 7.4.** Borde ideal, utilizado para determinar los límites de aplicabilidad del criterio de registro. (a) Gradiente orientado a  $0^\circ$  y (b) a  $45^\circ$ .



**Figura 7.5.** Incrementos de energía  $\overline{\Delta E}(r)$  para el borde ideal. (a) Curva para el borde orientado a  $0^\circ$ . (b) Curva para el borde orientado a  $45^\circ$ .

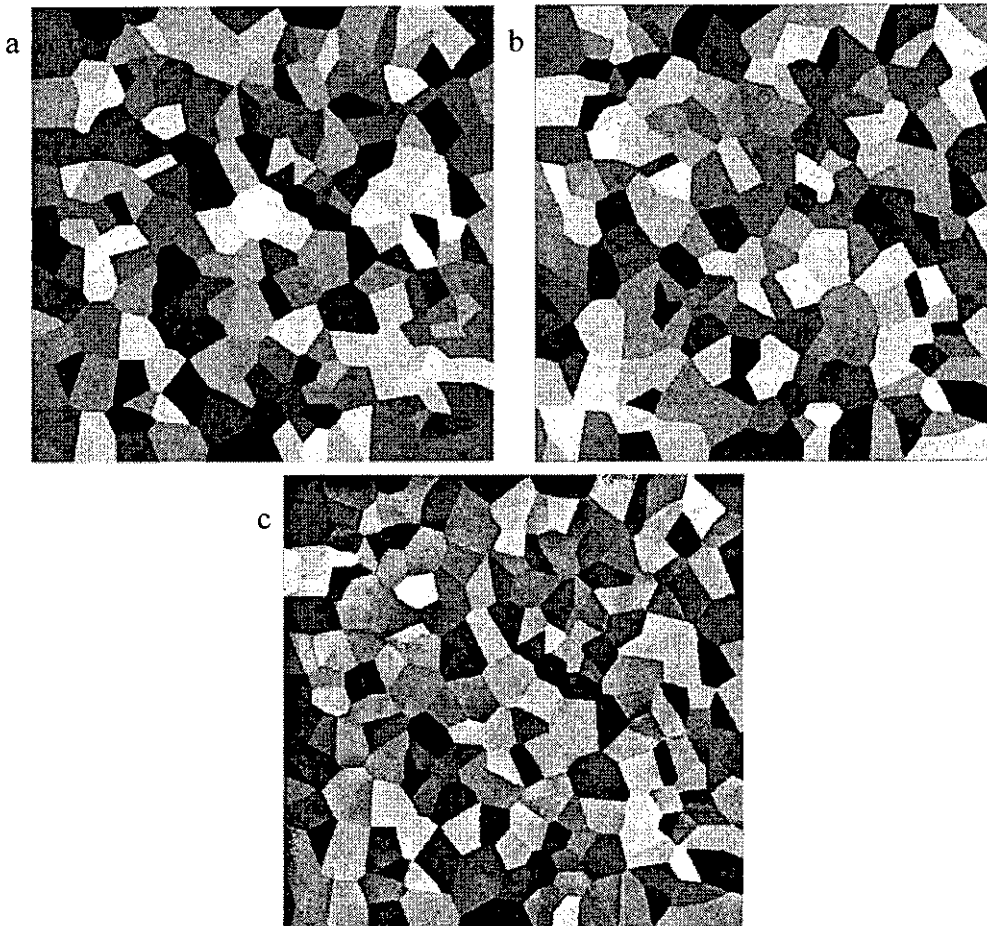
En la curva de la figura 7.5a se observa un crecimiento prácticamente lineal del incremento de energía en función de la distancia. Para el borde a  $45^\circ$  el crecimiento es menos uniforme, pero sigue siendo estrictamente creciente, lo cual reafirma la validez del criterio de evaluación. El máximo no se alcanza exactamente a 4 píxeles, debido a la anchura de banda radial,  $\alpha$ , y a las frecuencias accesorias que incluye la máscara  $M_{7 \times 7}$ .



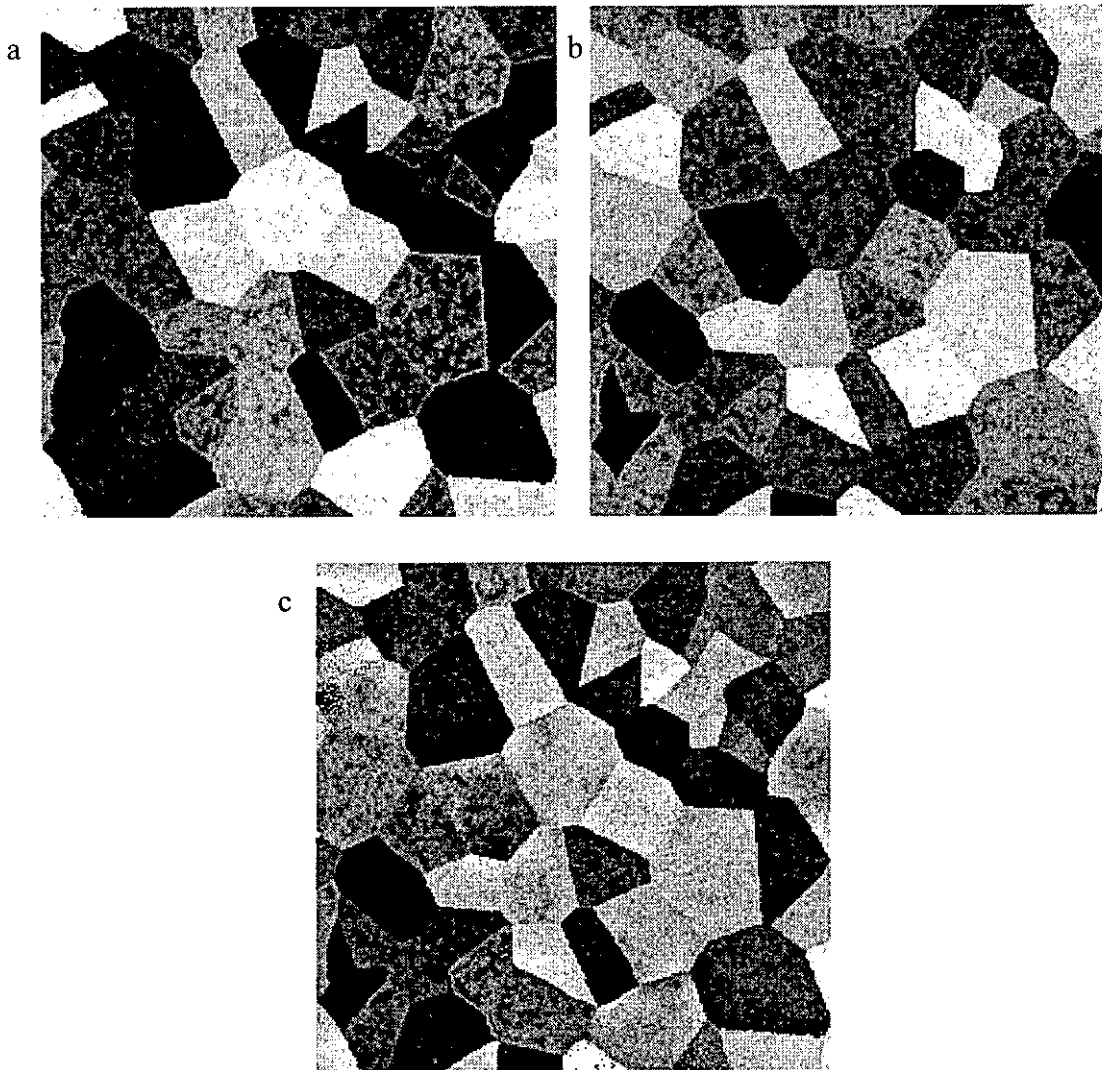
### 7.1.3.2. Imágenes sintéticas

Para validar el criterio desarrollado, se ha creado una pareja de imágenes sintéticas de modo que presente una distribución homogénea de segmentos en toda su superficie. Se ha utilizado como base el diagrama de Voronoi [64-68] de una distribución aleatoria de puntos, generado mediante nuestra técnica de crecimiento hexadecagonal, descrita en el artículo de Coiras, Santamaría y Miravet [49] y en el apéndice C.

A partir del diagrama se han creado dos imágenes diferentes (de 512x512 píxeles) asignando niveles de gris aleatorios a las regiones del diagrama y añadiendo después ruido impulsivo (de 40 niveles de gris de amplitud, al 50% de sus píxeles). Estas imágenes sintéticas, y el resultado de su fusión, se muestran en la figura 7.6. En el cálculo de los incrementos de energía se emplean únicamente las zonas centrales de las imágenes (de 256x256 píxeles, mostradas en la figura 7.7), para evitar efectos de borde al deformar la imagen fuente.



**Figura 7.6.** (a), (b) Par de imágenes sintéticas (512x512 píxeles) generadas a partir del diagrama de Voronoi de una distribución de puntos aleatoria. (c) Resultado de su fusión.

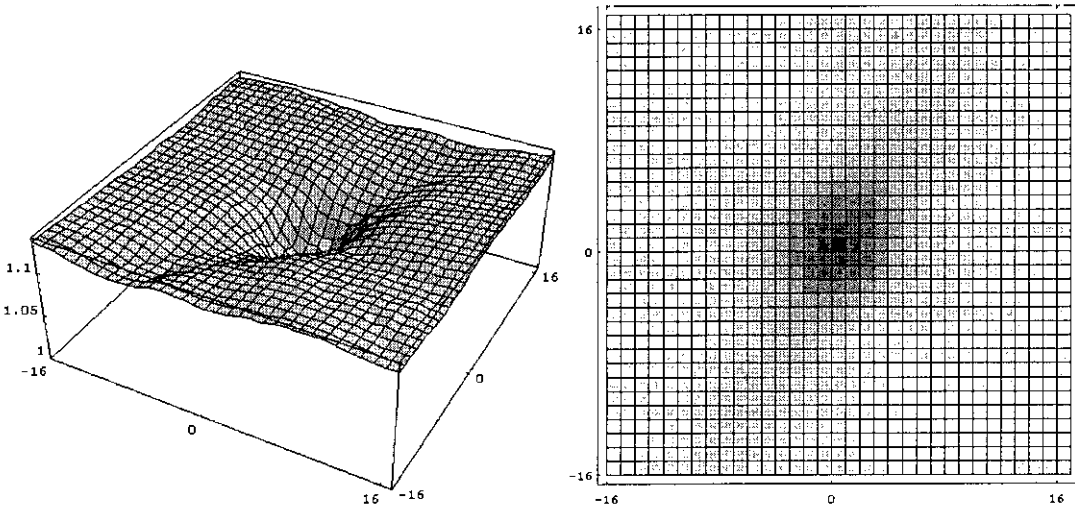


**Figura 7.7.** (a), (b) Regiones centrales (de 256x256 píxeles) de las imágenes sintéticas de la figura 7.6, utilizadas para evitar efectos de borde. (c) Fusión de (a) y (b).

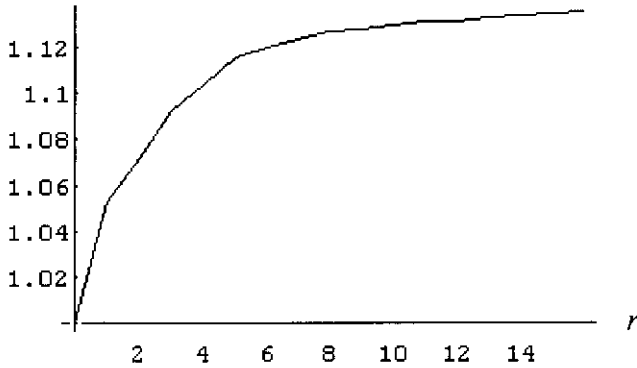
El mapa de distancias asociado a la traslación de la imagen 7.7a sobre la 7.7b se muestra en la figura 7.8. Como puede observarse, el mínimo de energía aparece muy localizado y la forma del mapa alrededor suyo presenta una distribución con simetría aproximadamente radial. Esto se debe al hecho de que la distribución de segmentos sobre la imagen es muy homogénea y sus orientaciones cubren también de forma homogénea todo el rango angular. Las imágenes reales suelen presentar más irregularidades.

La curva de distancias equivalentes obtenida a partir del mapa de la figura 7.8 puede verse en la figura 7.9. En este caso, al contrario que ocurriría con la imagen del borde ideal, el valor máximo sólo se alcanzará cuando no exista ningún solapamiento entre los bordes del par de imágenes, lo cual sólo podría suceder si se aplicase un desplazamiento del orden del

tamaño de las imágenes (256 píxeles). Los valores de la curva a partir de los 4 píxeles pueden oscilar, en función de la cantidad de bordes que se superpongan para cada desplazamiento. En cualquier caso, no se puede emplear el filtro  $M_{7 \times 7}$  para medir deformaciones equivalentes de más de cuatro píxeles, por lo que la obtención de desplazamientos equivalentes mayores, en función del incremento de energía asociado a un par de imágenes registradas, significan que las imágenes están mal registradas, aunque no se pueda determinar con exactitud el valor del desplazamiento equivalente.

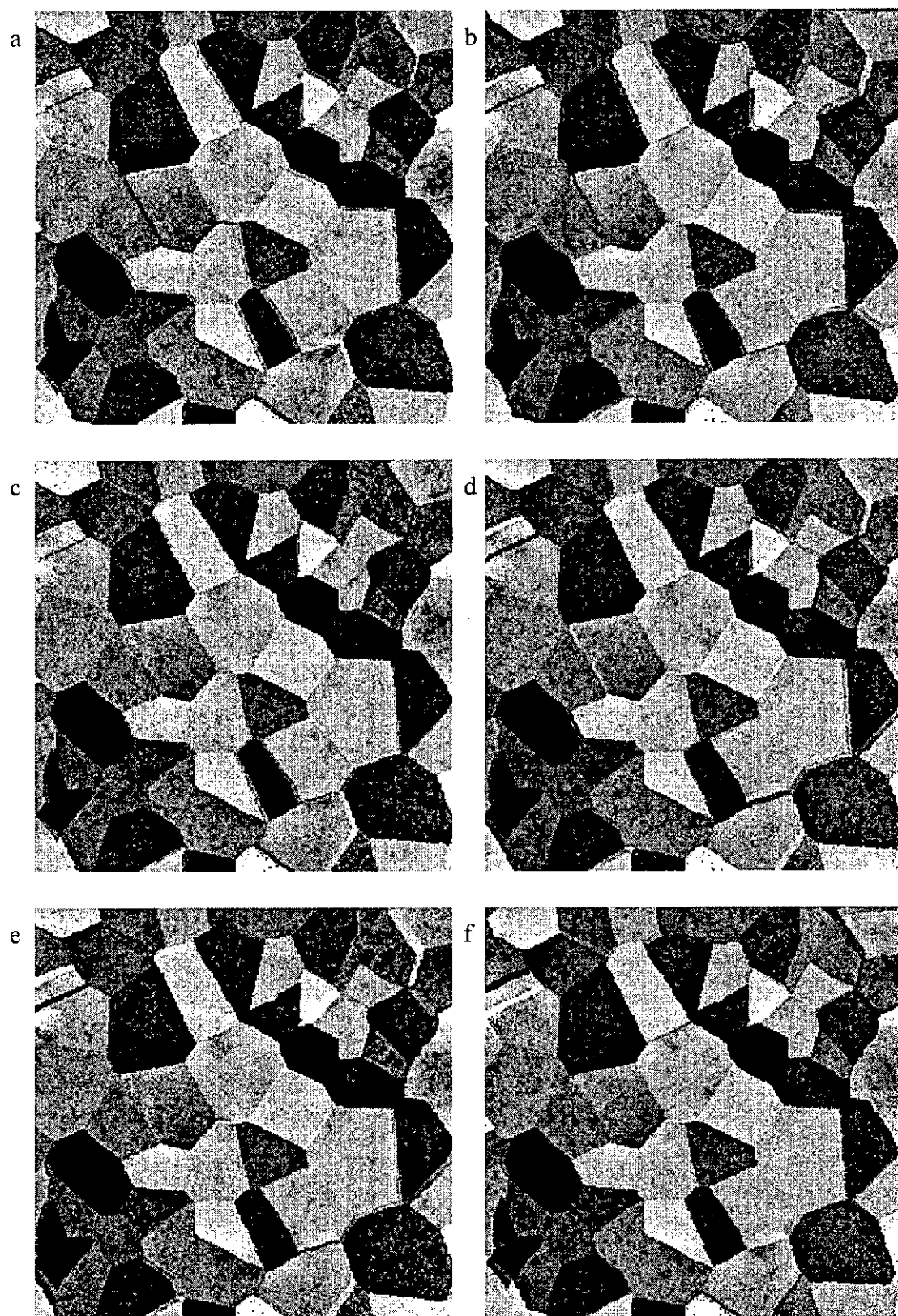


**Figura 7.8.** Mapa de incrementos de energía asociados a la traslación de la imagen sintética 7.7a sobre la 7.7b.



**Figura 7.9.** Incremento de energía  $\overline{\Delta E}(r)$  en función de la distancia para el par sintético de la figura 7.7, obtenido a partir del mapa de incrementos de la figura 7.8.

Para comprobar el funcionamiento del criterio de evaluación del registro se ha aplicado una serie de deformaciones de prueba. Para cada deformación se ha obtenido la imagen fundida y se ha calculado la distancia de desregistro equivalente. En la tabla 7.2 se muestran los resultados obtenidos.



**Figura 7.10.** Fusión del par de imágenes sintético después de aplicar las deformaciones de prueba: (a) traslación de (2, -2) píxeles, (b) rotación de  $2^\circ$ , (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%.

El efecto de la traslación en el resultado de la fusión de las dos imágenes sintéticas puede verse en la figura 7.10a, donde el desplazamiento ha sido de 2 píxeles en dirección horizontal y -2 píxeles en dirección vertical. El módulo de la distancia de traslación es, por tanto, 2.83 píxeles, mientras que la distancia equivalente obtenida a partir del incremento de energía en la banda de altas frecuencias es de 1.89 píxeles. La diferencia con la medida real significa que existen determinadas traslaciones de menor módulo que causan el mismo incremento de energía que la transformación particular elegida.

En la tabla 7.2 puede verse que las deformaciones más agresivas son las de rotación, barrilete y traslación, frente a las de escalado, sesgo y corsé, que causan un incremento de energía menor.

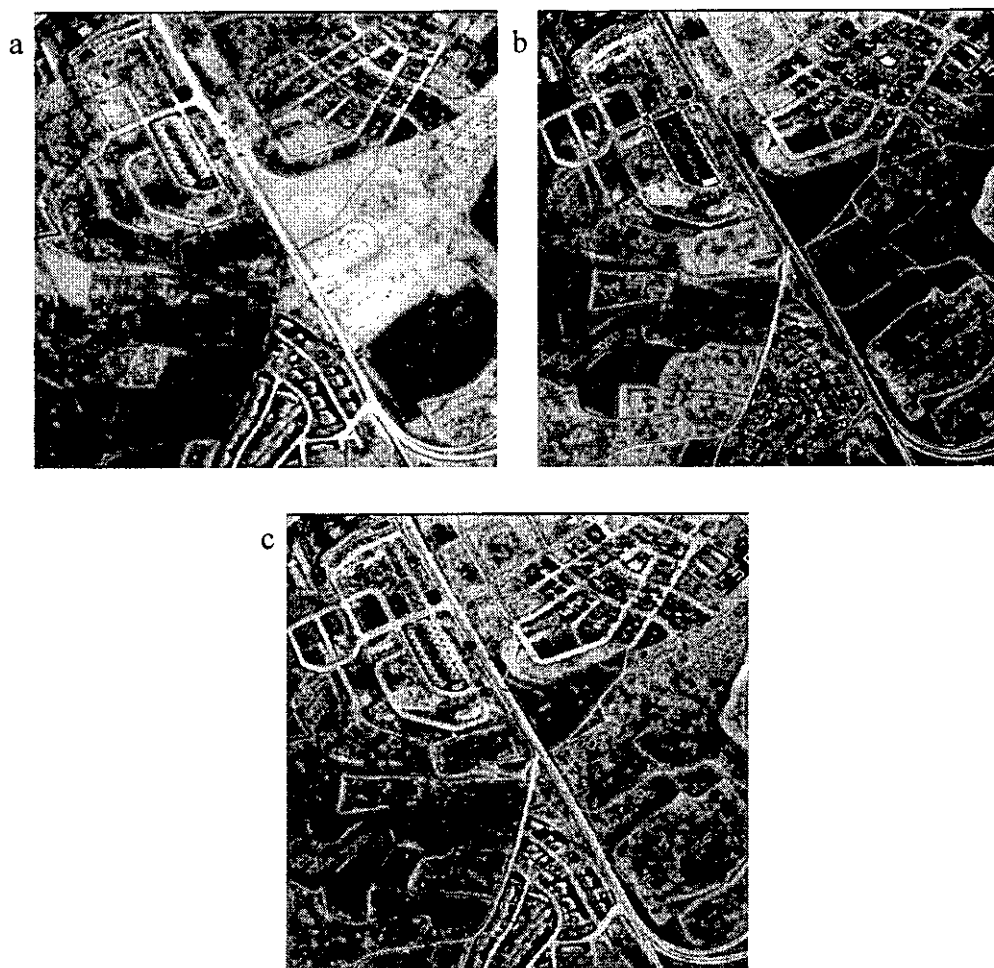
Deformación aplicada	$\bar{r}_T$
traslación (2, -2)	1.89
rotación 2°	2.33
sesgo 0.03	1.11
escalado 103%	1.56
distorsión en corsé 3%	0.97
distorsión en barrilete 3%	2.14

**Tabla 7.2.** Desplazamientos equivalentes (en píxeles) para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes sintéticas.

7.1.3.3. Imágenes Daedalus

El segundo par de prueba (figura 7.11) son dos imágenes aéreas de Navalcarnero, proporcionadas por el Instituto Nacional de Técnica Aeroespacial (INTA). Las imágenes, de 512x512 píxeles, han sido adquiridas mediante un sensor Daedalus DS-1268, y corresponden a las bandas espectrales de 2.08-2.35  $\mu\text{m}$  y 0.45-0.52  $\mu\text{m}$ . La pareja de imágenes, registrada por el propio sensor, presenta una distribución de información bastante homogénea.

También en este caso se han empleado únicamente la regiones centrales de las imágenes, que se muestran en la figura 7.12, para evitar efectos de borde.



**Figura 7.11.** Bandas (a) de 2.08 a 2.35  $\mu\text{m}$  y (b) de 0.45 a 0.52  $\mu\text{m}$  de una imagen Daedalus de Navalcarnero. (c) El resultado de su fusión.

El mapa de incrementos y la distribución radial de energía se muestran en las figuras 7.13 y 7.14, respectivamente. En este caso, el mapa de incrementos no presenta una distribución tan isótropa como el obtenido a partir de las imágenes sintéticas, aunque la distribución en las cercanías del mínimo aún posee una simetría aproximadamente radial.

Los resultados de la aplicación del procedimiento de evaluación del registro para las regiones centrales de las imágenes Daedalus, en base a las mismas deformaciones que fueron aplicadas al par sintético, se muestran en la tabla 7.3. De nuevo las deformaciones que mayor incremento de energía provocan son las de rotación, barrilete y traslación, y las que menos afectan son las de sesgo, escalado y corsé.

En la figura 7.15 pueden verse los resultados de la fusión de la imagen fuente distorsionada y la imagen destino para cada una de las deformaciones de prueba.

Deformación aplicada	$\bar{r}_T$
traslación (2, -2)	1.71
rotación 2°	2.07
sesgo 0.03	0.78
escalado 103%	0.54
distorsión en corsé 3%	0.31
distorsión en barrilete 3%	2.81

**Tabla 7.3.** Desplazamientos equivalentes (en píxeles) para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes Daedalus.



**Figura 7.12.** (a), (b) Regiones centrales (de 256x256 píxeles) de las imágenes Daedalus de la figura 7.11, utilizadas para evitar efectos de borde. (c) y su fusión.



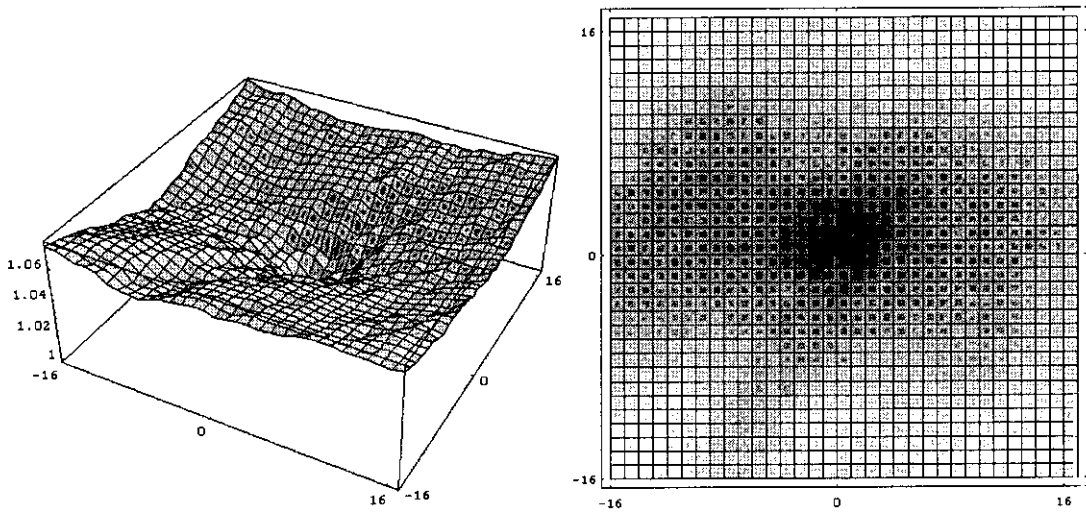


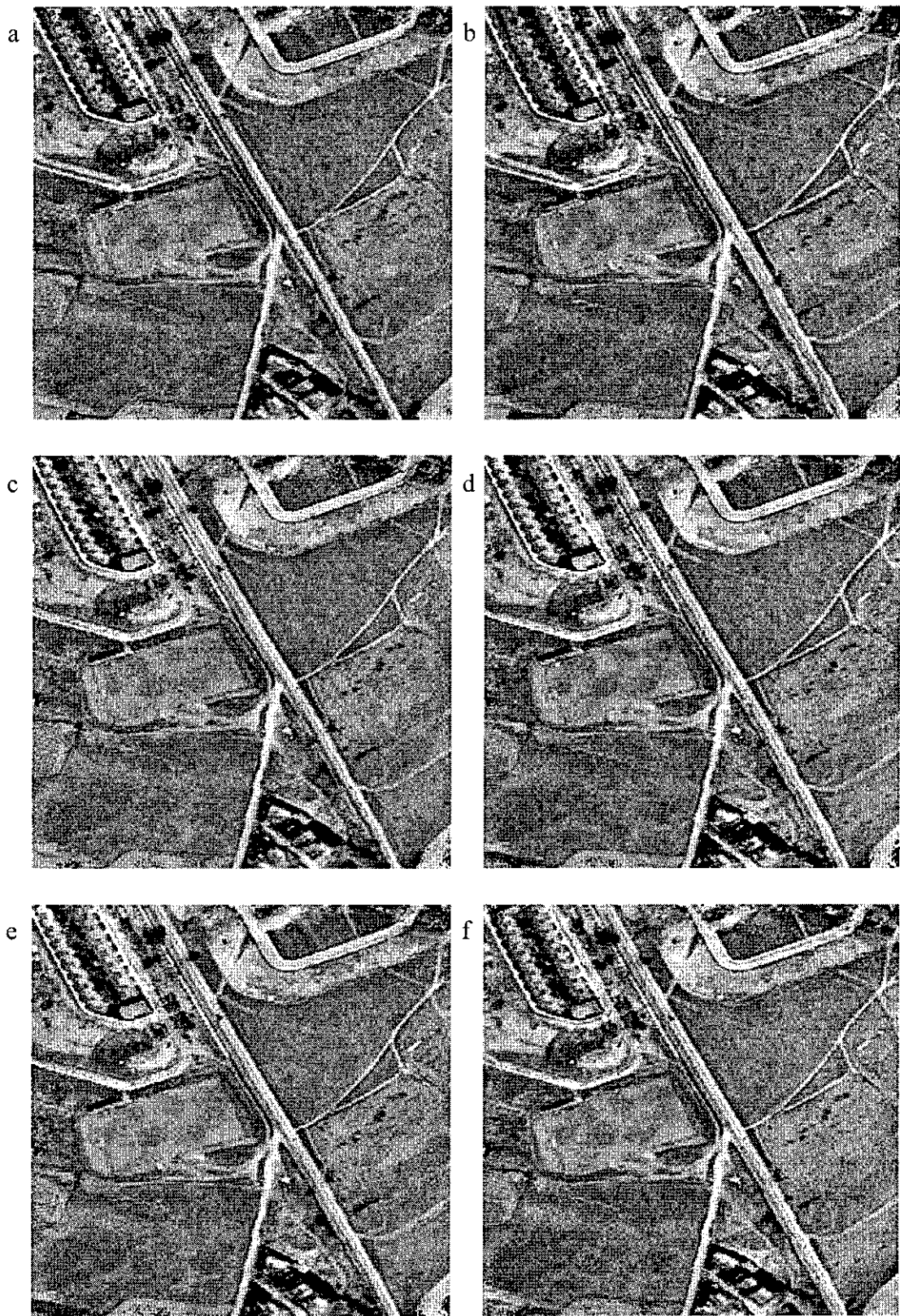
Figura 7.13. Mapa de incrementos de energía del par de imágenes Daedalus de la figura 7.12.



Figura 7.14. Incremento de energía  $\Delta E(r)$  en función de la distancia para las imágenes de la figura 7.12, obtenido a partir del mapa de incrementos de la figura 7.13.

Como puede observarse, los desregistros en las imágenes Daedalus fundidas son más difíciles de apreciar a simple vista. En el par sintético, la relativa limpieza de las fronteras entre regiones hace que los bordes dobles destaquen claramente. En la pareja de imágenes reales ya existe cierta duplicidad de contornos en los objetos (todas las carreteras y caminos, por ejemplo, definen dos bordes paralelos), con lo que el efecto del desregistro sobre la fusión es más difícil de percibir. El criterio de evaluación del desregistro, sin embargo, proporciona desplazamientos equivalentes similares a los obtenidos para el par sintético, lo que reafirma su validez como método para la evaluación de la calidad de registro para cualquier tipo de imágenes. Las medidas, obviamente, no son exactamente iguales, ya que dependen del modo en el que se distribuye la información en los bordes, que es diferente para cada par de imágenes.





**Figura 7.15.** Fusión del par de imágenes Daedalus después de aplicar las deformaciones de prueba: (a) traslación de (2, -2) píxeles, (b) rotación de 2°, (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%.

## 7.2. Evaluación basada en la determinación del campo vectorial de desplazamientos locales

El criterio de evaluación del registro basado en la determinación del incremento de energía de altas frecuencias proporciona una valoración global “media” del desregistro de las imágenes, en función de la degradación del aspecto visual de la imagen fundida (esto es, depende del contenido de las propias imágenes). En algunas ocasiones, sin embargo, puede ser interesante precisar el nivel de desplazamiento real de determinados puntos o zonas de las imágenes, sin tener en cuenta la cantidad de información que contengan. Es el caso, por ejemplo, de la utilización del registro para la calibración de cámaras u otros dispositivos de adquisición de imágenes. Para ello, y puesto que normalmente se desconoce la función de deformación,  $T_{opt}^{-1}$ , es necesario desarrollar algún método para evaluar el campo vectorial de desplazamientos locales a partir de la imágenes del par de registro.

Al igual que en la evaluación basada en Gabor, se necesita disponer de la imagen original,  $T_{opt}(I_I)$ , que, salvo en los casos de prueba, no se conoce. Por ello, realizaremos también aquí un registro manual para obtener una aproximación de la imagen original. El campo puede calcularse entonces mediante la correlación [14-17] de pequeñas ventanas, aunque también es posible emplear técnicas más elaboradas para esta tarea, como las basadas en la evaluación del flujo óptico [21-24].

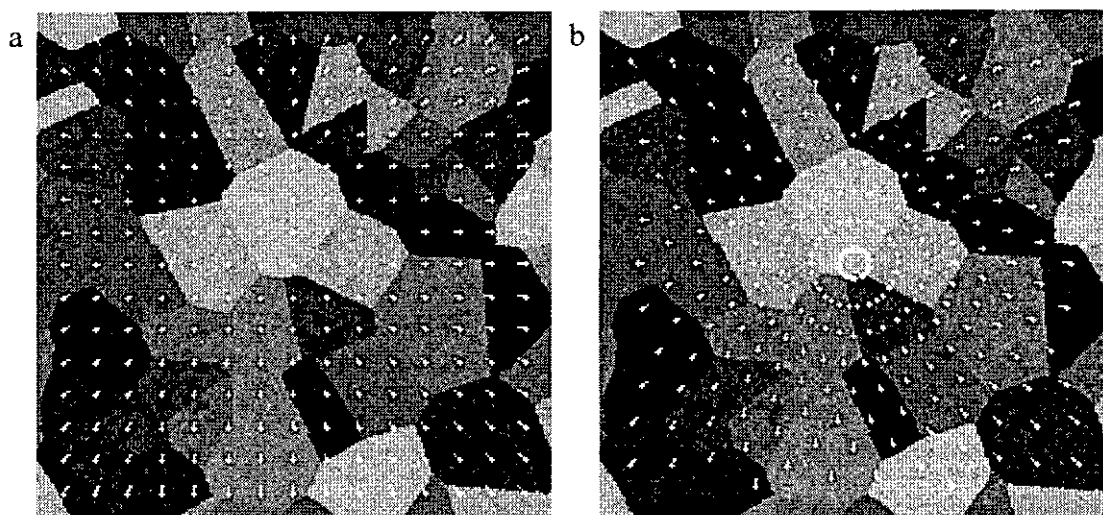
Para nuestro desarrollo hemos empleado el método para el cálculo del flujo óptico utilizado por Fuh y Maragos [22], principalmente por su facilidad de implementación. El método se basa en la creación de una batería de ventanas, obtenidas mediante transformaciones afines de la ventana fuente, que son posteriormente correlacionadas sobre una región determinada de la imagen destino para obtener el vector de traslación del centro de la ventana. A modo de ejemplo, se muestran en la figura 7.16 los resultados de la aplicación de esta técnica de determinación del campo de desplazamientos para una de las deformaciones de prueba, empleando dos formas distintas de distribuir los puntos en los que se evalúa el campo.

### 7.2.1. Obtención de un valor de desplazamiento medio equivalente

Para representar visualmente el campo vectorial asociado al desregistro basta con calcular los vectores de desplazamiento en una rejilla cuadrada de puntos equiespaciados, como la

que se muestra en la figura 7.16a. A partir de ella, se puede extraer directamente la media del módulo de los vectores del campo de desplazamientos y utilizarla como valor medio del desplazamiento de desregistro. Sin embargo, si se desea obtener una medida de desplazamiento medio similar a la que resulta de la aplicación del método de evaluación basado en Gabor, no es éste el modo más conveniente de hacerlo.

Efectivamente, la utilización de una rejilla cuadrada valora de la misma forma los puntos situados en la región central de la imagen (que es la que más importancia visual tiene) y los de regiones cercanas al contorno de la imagen (de menor relevancia y generalmente con deformaciones muy superiores a las de la región central). Es preferible, por tanto, emplear otro tipo de distribuciones que valoren más la región central de la imagen, como, por ejemplo, tomando un mismo número de muestras,  $N_s$ , en anillos equiespaciados desde el centro de la imagen. Esta configuración se muestra en la figura 7.16b, donde puede apreciarse la inclusión de una rotación adicional de  $\pi/N_s$  en los anillos impares, para recubrir el espacio imagen de forma más homogénea.



**Figura 7.16.** Dos formas de distribuir puntos en una imagen para la evaluación del valor de desplazamiento medio equivalente. (a) Rejilla cuadrada y (b) malla distribuida radialmente, con el mismo número de puntos de muestra en cada anillo.

A la hora de determinar el vector de desplazamiento local de un punto de la imagen mediante el empleo de cualquier método basado en correlación, conviene tener en cuenta si la cantidad de información presente en sus inmediaciones es suficiente para garantizar la calidad del máximo de correlación que se obtenga. Para la validación de las ventanas de correlación empleadas, hemos utilizado el operador de interés descrito en el libro de Haralick y Shapiro [10]. Este operador utiliza los bordes extraídos en la ventana para

calcular una *elipse de confianza* (cuya forma y orientación determinan la existencia de una posible dirección privilegiada) y un valor o peso que depende de la cantidad de información presente en la zona. Si no hay suficiente información, no es posible obtener un valor de correlación fiable en la zona, y si existe una orientación privilegiada muy acusada, aparecerá un problema de apertura en la evaluación de desplazamientos locales orientados en dicha dirección.

Para la determinación de la elipse de confianza de una ventana  $W$ , contenida en una imagen  $I$ , debe calcularse la matriz normal siguiente:

$$\mathbf{N} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} = \begin{pmatrix} \sum_{(x,y) \in W} \left( \frac{\partial I}{\partial x}(x,y) \right)^2 & \sum_{(x,y) \in W} \frac{\partial I}{\partial x}(x,y) \cdot \frac{\partial I}{\partial y}(x,y) \\ \sum_{(x,y) \in W} \frac{\partial I}{\partial x}(x,y) \cdot \frac{\partial I}{\partial y}(x,y) & \sum_{(x,y) \in W} \left( \frac{\partial I}{\partial y}(x,y) \right)^2 \end{pmatrix} \quad (7.17)$$

donde  $I(x, y)$  es la intensidad (nivel de gris) de cada punto de la imagen. Para la determinación de los gradientes hemos utilizado el método de Sobel [9-13], aunque en principio la elección del extractor de bordes no es crítica.

En función de la matriz normal  $\mathbf{N}$ , se determinan el peso o cantidad de información,  $w$ , de la ventana, la dirección  $\phi$  del eje mayor de la elipse de confianza y su factor de forma  $q$  (igual a uno menos la excentricidad), mediante las fórmulas siguientes:

$$w = 2 \frac{\det(\mathbf{N})}{\text{tr}(\mathbf{N})} \quad \phi = \frac{1}{2} \text{tg}^{-1} \left( \frac{-2N_{12}}{N_{11} - N_{22}} \right) \quad q = 4 \frac{\det(\mathbf{N})}{\text{tr}^2(\mathbf{N})} \quad (7.18)$$

donde  $\det(\mathbf{N})$  y  $\text{tr}(\mathbf{N})$  indican el determinante y la traza, respectivamente, de la matriz normal de la ventana  $W$  (7.17).

En nuestra implementación particular, se parte de una malla de puntos,  $M = \{p_i\}$ , distribuidos de una determinada forma (rejilla cuadrada, radial, etc) sobre la imagen fuente correctamente registrada,  $T_{opt}^{-1}(I_I)$ . Centrada en cada punto  $p_i$  de la malla se extrae una ventana que será posteriormente buscada en la imagen fuente  $I_I$  (bien registrada, si deseamos evaluar la calidad del método de registro empleado, o bien sin registrar, si lo que se pretende es obtener una representación del campo vectorial de desregistro). Se calculan entonces los parámetros de la elipse de confianza (7.18) para cada una de estas ventanas, y

aquéllas para las que el peso  $w$  sea inferior a 0.05, o que presenten un factor de forma  $q$  inferior a 0.2 son descartadas.

Las ventanas que sean consideradas válidas según este criterio se buscan en la imagen fuente  $I_I$  mediante el método de Fuh y Maragos [22], y se obtiene el conjunto  $M' = \{p_i'\}$  de puntos  $p_i'$  sobre  $I_I$  correspondientes a los puntos  $p_i$  de la malla  $M$  sobre  $T_{opt}^{-1}(I_I)$ . El conjunto de vectores de desplazamientos locales evaluado en los puntos de la malla es, por tanto:

$$V = \{v_i = (v_i^x, v_i^y) = (p_i' - p_i) \mid p_i \in M, p_i' \in M'\}$$

Para la estimación de un valor de desplazamiento medio equivalente de la imagen  $I_I$  respecto de la imagen perfectamente reconstruida,  $T_{opt}^{-1}(I_I)$ , utilizamos una malla radial (figura 7.16b) y asignamos un peso,  $\chi$ , a cada uno de los vectores  $v_i$  del campo en función de la forma  $q$  y la orientación  $\phi$  de la elipse de confianza de la ventana  $W(v_i)$ , centrada en el punto inicial del vector  $v_i$  (es decir,  $p_i \in M$ ), de forma que sean menos valorados aquellos vectores con una orientación cercana a la dirección del eje principal de la elipse cuando ésta tenga una excentricidad acusada. La función de peso que hemos elegido a tal fin es la siguiente:

$$\chi(v_i) = q + (1 - q) \cdot |\sin(\alpha)| \quad (7.19)$$

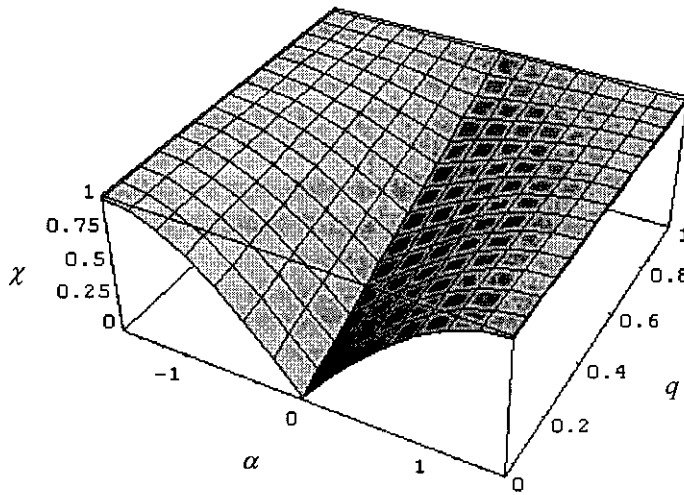
donde  $\alpha$  es la diferencia de orientaciones entre el vector  $v_i$  y el eje principal de la elipse de confianza:

$$\alpha \equiv \phi - \text{tg}^{-1}\left(\frac{v_i^x}{v_i^y}\right) \quad (7.20)$$

Como puede verse en la figura 7.17, cuando el vector está orientado perpendicularmente a la dirección del eje principal de la elipse de confianza ( $\alpha = k\pi/2$ ,  $k \in \mathbf{Z}$ ), o cuando la elipse es una circunferencia ( $q = 1$ ), la función de peso  $\chi$  vale 1. Únicamente cuando la dirección del vector se aproxime a la del eje principal de la elipse existirá un potencial problema de apertura, que será más acusado cuanto más se acerque  $q$  a cero. Esto se refleja en la función de peso,  $\chi$ , que vale exactamente  $q$  cuando dichas orientaciones coinciden ( $\alpha = 0$ ).

Empleando la función de peso  $\chi$  y el conjunto de vectores  $V$  puede obtenerse entonces un valor de desplazamiento medio equivalente,  $\bar{r}(V)$ , similar al resultante de la aplicación del criterio de evaluación del registro basado en Gabor:

$$\bar{r}(V) = \frac{\sum_{v_i \in V} \chi(v_i) \cdot \|v_i\|}{\text{Card}(V)} \quad (7.21)$$



**Figura 7.17.** Valor del peso  $\chi$  en función del factor de forma  $q$  y de la diferencia de orientación  $\alpha$  entre un vector de desplazamiento y la dirección del eje principal de la elipse de confianza.

A modo de ejemplo, se muestran a continuación (figuras 7.18 y 7.19) los campos vectoriales calculados mediante este método para las mismas imágenes de prueba utilizadas en el apartado 7.2. Los valores de desplazamientos medios obtenidos, para las imágenes de prueba sintéticas y Daedalus, se presentan en las tablas 7.4 y 7.5, respectivamente.

Como puede observarse, las medidas del valor medio del módulo de los vectores del campo son muy similares en las dos tablas 7.4 y 7.5 (deben serlo, puesto que las funciones de deformación aplicadas son exactamente las mismas). Las pequeñas diferencias existentes se deben a la eliminación de algunos de los puntos de la rejilla en función de sus elipses de confianza. En cuanto al desplazamiento medio equivalente, las variaciones, además de depender de los puntos descartados, están influenciadas por el contenido de bordes de las imágenes; y lo están débilmente, puesto que los puntos para los que la elipse de confianza es muy excéntrica son eliminados en la fase de filtrado.

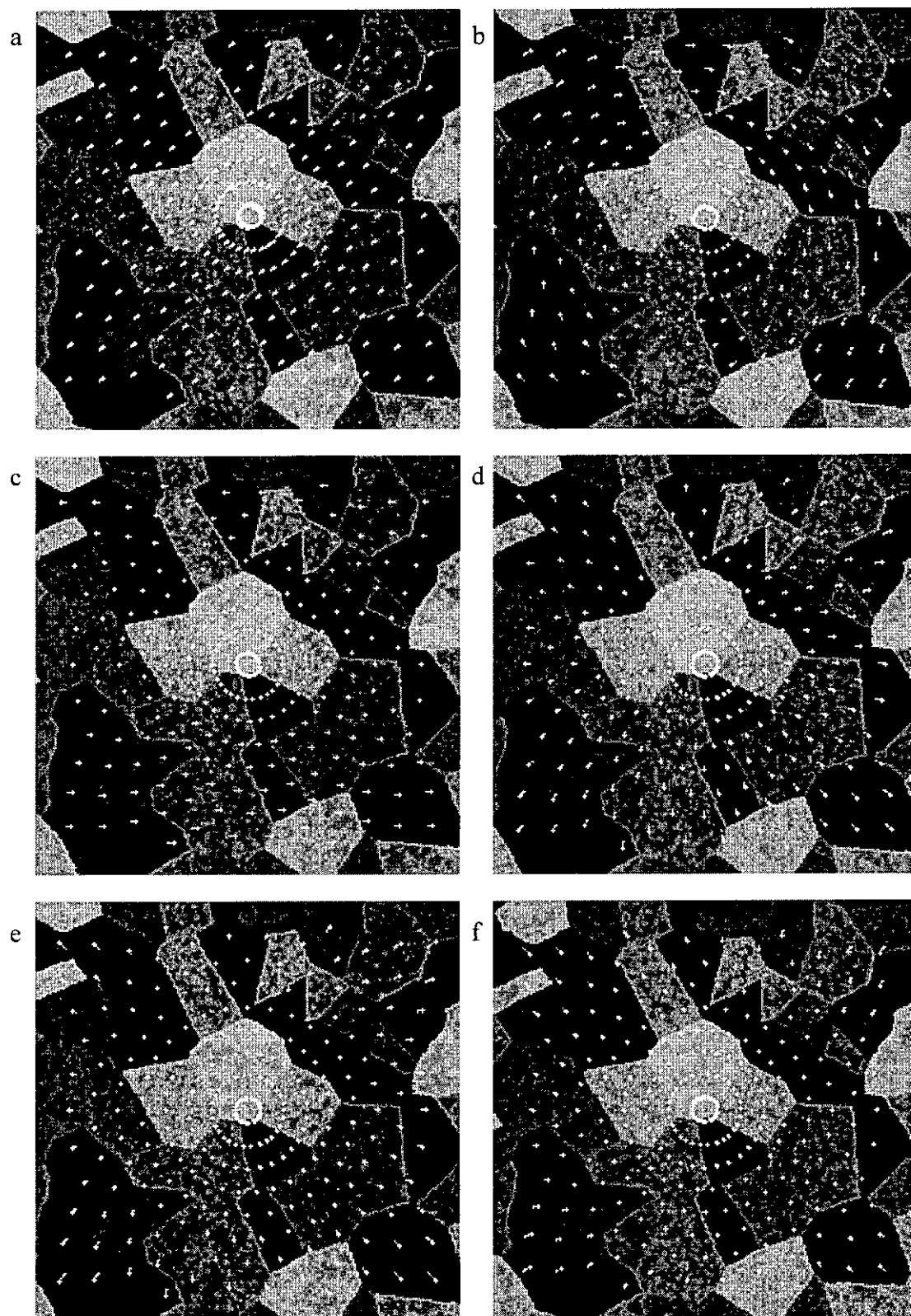
Deformación aplicada	$\ V\ $	$\bar{r}(V)$
traslación (2, -2)	2.83	2.63
rotación 2°	2.35	2.20
sesgo 0.03	1.27	1.10
escalado 103%	2.05	1.88
distorsión en corsé del 3%	1.50	1.38
distorsión en barrilete del 3%	1.25	1.16

**Tabla 7.4.** Módulos medios y desplazamientos medios equivalentes (ambos en píxeles), calculados en función del campo vectorial de desregistro, para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes sintéticas.

Deformación aplicada	$\ V\ $	$\bar{r}(V)$
traslación (2, -2)	2.83	2.68
rotación 2°	2.40	2.18
sesgo 0.03	1.26	1.05
escalado 103%	2.06	1.82
distorsión en corsé del 3%	1.51	1.34
distorsión en barrilete del 3%	1.24	1.08

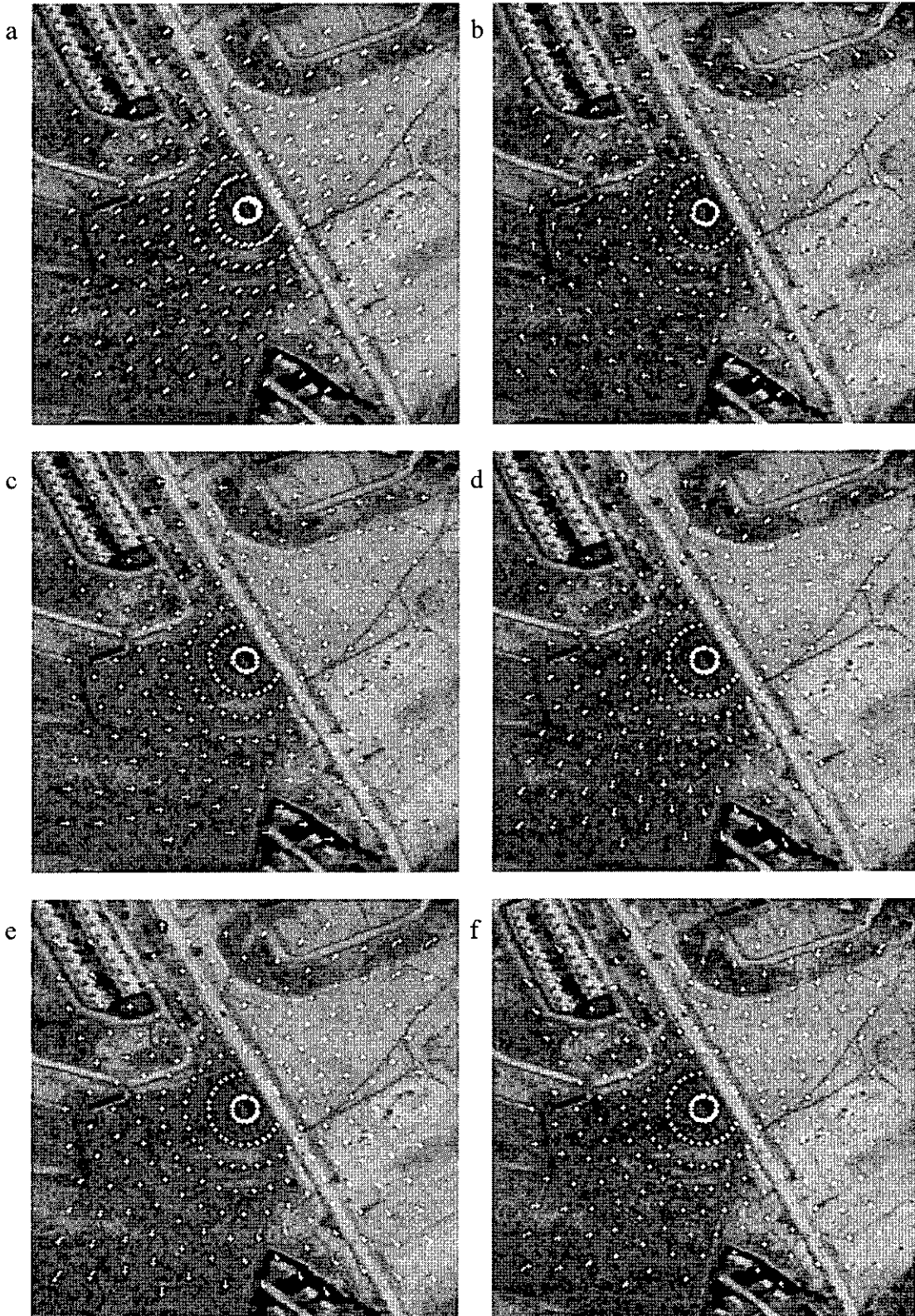
**Tabla 7.5.** Módulos medios y desplazamientos medios equivalentes (ambos en píxeles), calculados en función del campo vectorial de desregistro, para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes Daedalus.

A modo comparativo pueden contrastarse estos resultados con los obtenidos mediante el criterio de evaluación basado en Gabor, recogidos en las tablas 7.2 y 7.3. Las diferencias son notables, debido a la fuerte dependencia que presenta el criterio de Gabor respecto al contenido de la escena.



**Figura 7.18.** Campos vectoriales correspondientes a las funciones de deformaciones de prueba aplicadas a la imagen fuente sintética 7.5a. (a) traslación de (2, -2) píxeles, (b) rotación de  $2^\circ$ , (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%.





**Figura 7.19.** Campos vectoriales correspondientes a las funciones de deformación aplicadas a la imagen fuente Daedalus 7.10a. (a) traslación de (2, -2) píxeles, (b) rotación de  $2^\circ$ , (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%.

### 7.3. Ejemplo de aplicación de los criterios de evaluación desarrollados

A continuación se aplicarán los criterios de evaluación desarrollados para la valoración del registro de un par real de imágenes IR-visible. Las imágenes han sido registradas mediante el método de casamiento de conjuntos de segmentos lineales (capítulo 5), y la transformación de registro resultante ha sido posteriormente refinada (capítulo 6) para generar un polinomio de registro de tercer orden. En la figura 7.19 se muestran el par de registro, las imágenes fuente IR registradas mediante los dos modelos de transformación y la imagen de referencia, registrada manualmente con un modelo de tercer orden, para la evaluación de las imágenes registradas automáticamente. Visualmente, los resultados de las fusiones son muy parecidos (lo que reafirma la robustez de nuestro método de registro), aunque existen pequeñas diferencias entre las imágenes, que pueden ser cuantificadas mediante los criterios de evaluación del registro desarrollados. Estas diferencias se aprecian principalmente en la esfera de la farola, el portal y las ventanas del edificio de la derecha, y los extremos izquierdo e inferior de las imágenes IR. La mayor parte de los defectos no corregidos en la imagen registrada aún se deben a la imposibilidad de reconstruir las deformaciones de distorsión (apreciables en la curvatura de las paredes de los edificios en la imagen IR) y de perspectiva (ligeramente apreciables en el suelo de la parte inferior de la imagen, más cercana a las cámaras). La imagen restaurada mediante el modelo de tercer orden presenta una calidad visual equivalente a la imagen registrada de referencia, y únicamente se perciben artefactos en los contornos de la imagen IR reconstruida.

En la figura 7.20 se muestran la región de las imágenes considerada para los cálculos de la calidad de registro (7.20a) y las componentes de alta frecuencia de cada una de las imágenes, empleadas para el cálculo del desplazamiento medio equivalente basado en Gabor. La componente de la imagen visible (7.20b) contiene más bordes que las de las imágenes IR registradas (7.20c, d y e), lo que significa que la mayoría de los detalles (correspondientes a la banda de altas frecuencias espaciales) de las imágenes fundidas procederán de la primera. También puede apreciarse, no obstante, que existen estructuras que únicamente aparecen en la imagen IR, como el poste de la farola, la tapa metálica de la boca de riego sobre el césped o la sombra del portal sobre el cristal que lo rodea. En la componente de Gabor de las imágenes fundidas (7.20f, g y h) aparecen tanto las aportaciones de la imagen visible como las de las infrarrojas.

El mapa de incrementos de energía (figura 7.21, fila superior), aproximadamente con simetría radial alrededor del mínimo, presenta anisotropías en forma de punto de silla,

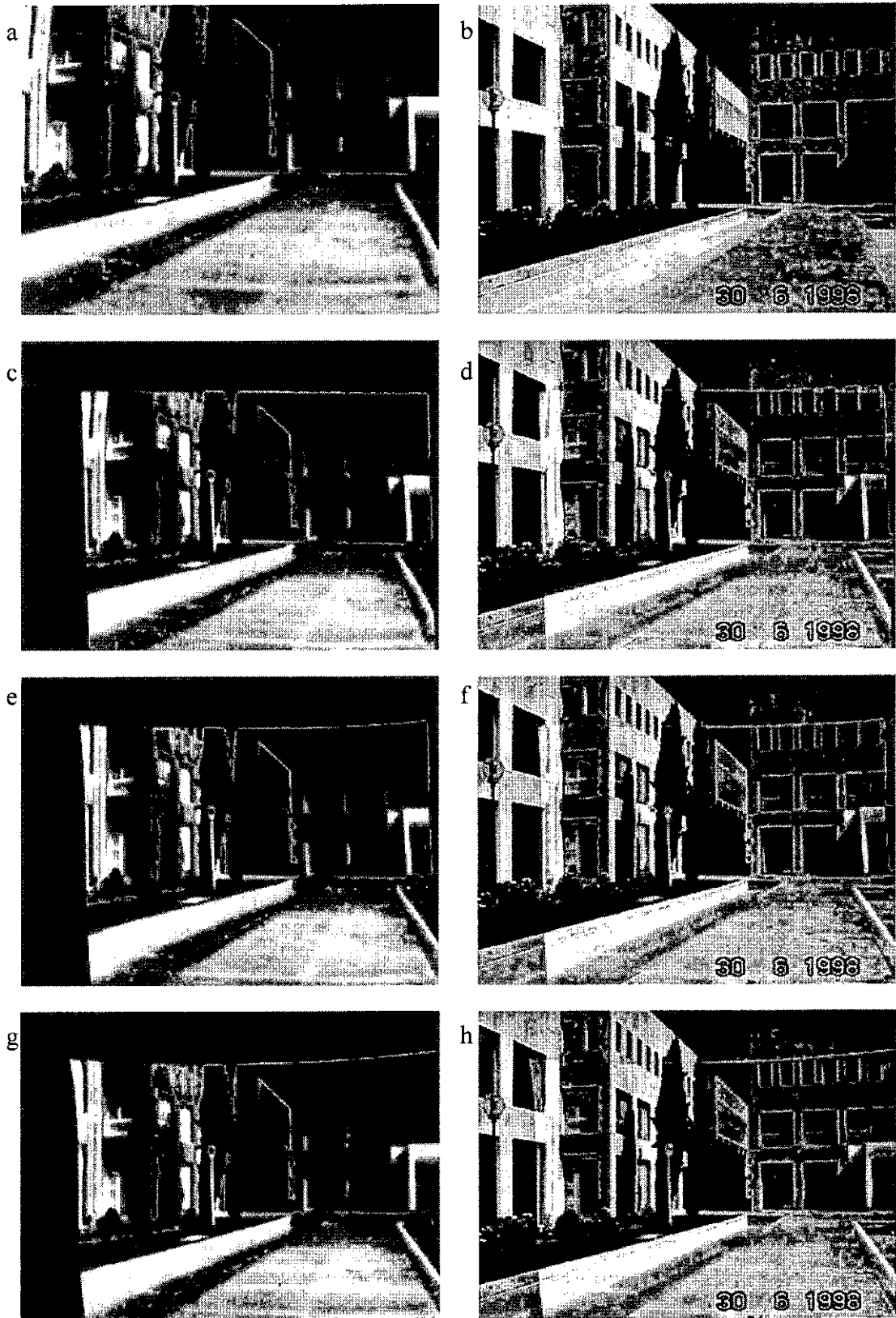
debido a la existencia de dos direcciones privilegiadas en los objetos que aparecen en la imagen. Estas dos direcciones (paralela y perpendicular al suelo) aparecen en la mayoría de las construcciones y objetos elaborados por el hombre, por lo que las escenas que contengan elementos de este tipo presentarán con frecuencia mapas de energía con esa forma particular.

La curva de incrementos de energía en función de la distancia al mínimo (figura 7.21, fila inferior) se obtiene a partir del mapa y se utiliza para la evaluación de las imágenes IR registradas automáticamente. La curva presenta una pendiente menor que las que se obtenían para las imágenes de prueba (figuras 7.9 y 7.14), lo que significa que la degradación en la imagen fundida acusa menos los efectos del desregistro que aquéllas. Esa menor pendiente revaloriza aún más los buenos resultados obtenidos (tabla 7.6).

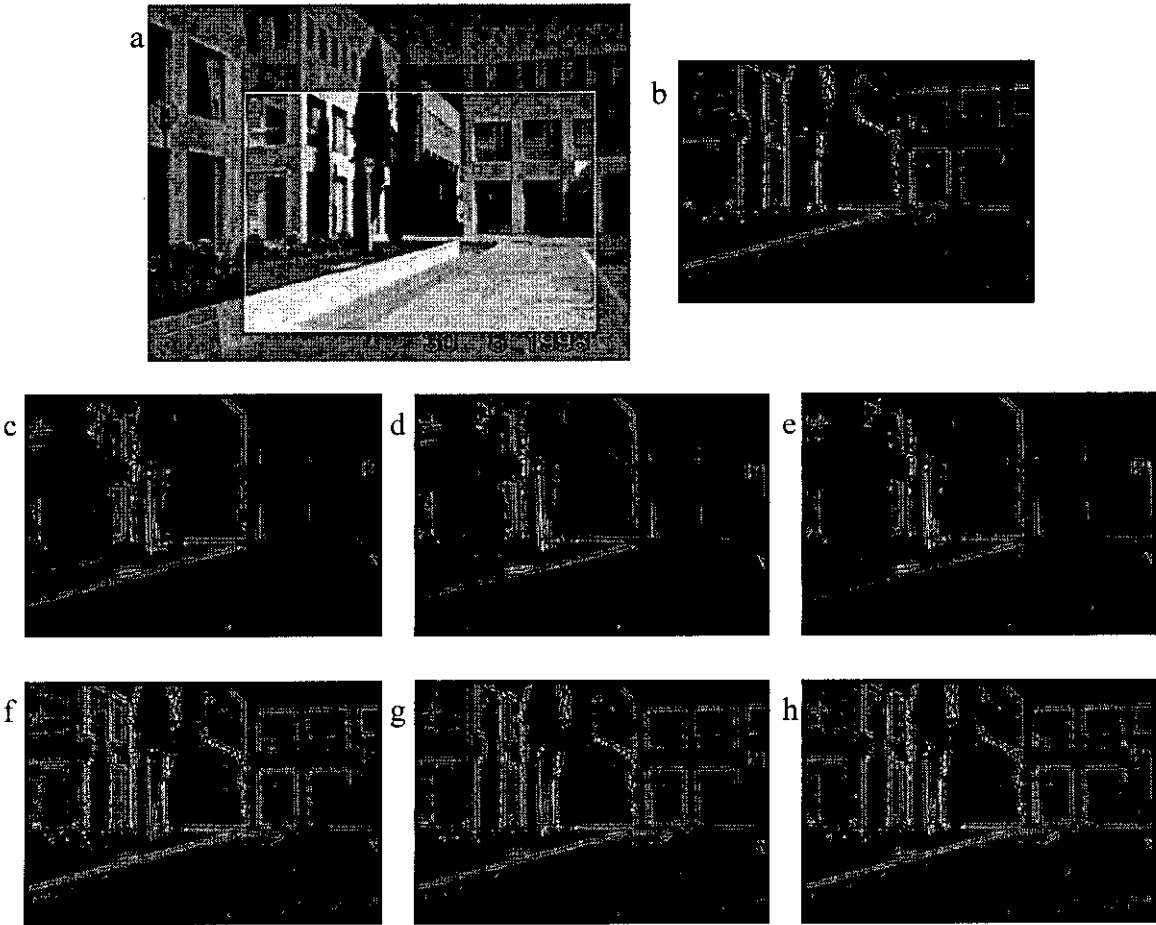
La evaluación mediante el criterio basado en el campo vectorial de desplazamientos locales aporta más información sobre las diferencias existentes entre las imágenes registradas de manera automática y la imagen de referencia, registrada manualmente. Como puede verse en la figura 7.22, las zonas con mayor cantidad de información presentan valores de desplazamiento menores (primer cuadrante, principalmente). En las regiones en las que existe una clara dirección privilegiada (zona izquierda, sentido vertical) la relevancia del módulo de los desplazamientos es menor, puesto que sus orientaciones son paralelas a dicha dirección. La existencia de esa dirección privilegiada en esa zona también explica porqué el método de registro ha sido menos preciso allí, puesto que, en el fondo, su funcionamiento se basa en los cruces de los segmentos detectados en las imágenes, más escasos en la zona. Tampoco son relevantes los desplazamientos existentes en la zona del suelo (inferior derecha), donde no existe prácticamente ninguna información significativa (como puede verificarse mediante la inspección de las componentes de Gabor, en la figura 7.20). Los resultados para este criterio se recogen también en la tabla 7.6.

		Criterio de Evaluación		
		Gabor	Campo Vectorial	
		$\bar{r}_T$	$\ \bar{V}\ $	$\bar{r}(V)$
Modelo	afín	1.22	2.16	1.94
	tercer orden	1.01	2.13	1.88

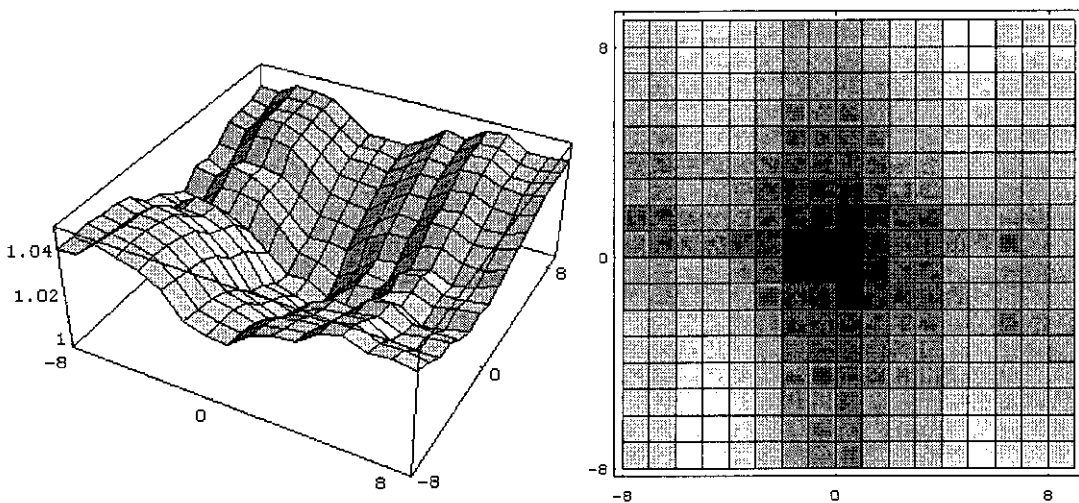
**Tabla 7.6.** Desplazamientos medios equivalentes (en píxeles), para los dos modelos de transformación en base a los criterios de evaluación desarrollados.



**Figura 7.20.** Imágenes IR registradas y resultados de su fusión con la imagen visible. (a) y (b) Par de registro IR-visible. (c) y (d) Registro y fusión para el modelo afín de transformación. (e) y (f) Registro y fusión para el modelo de tercer orden. (g) y (h) Resultados para la imagen de referencia óptima, generada mediante registro manual.



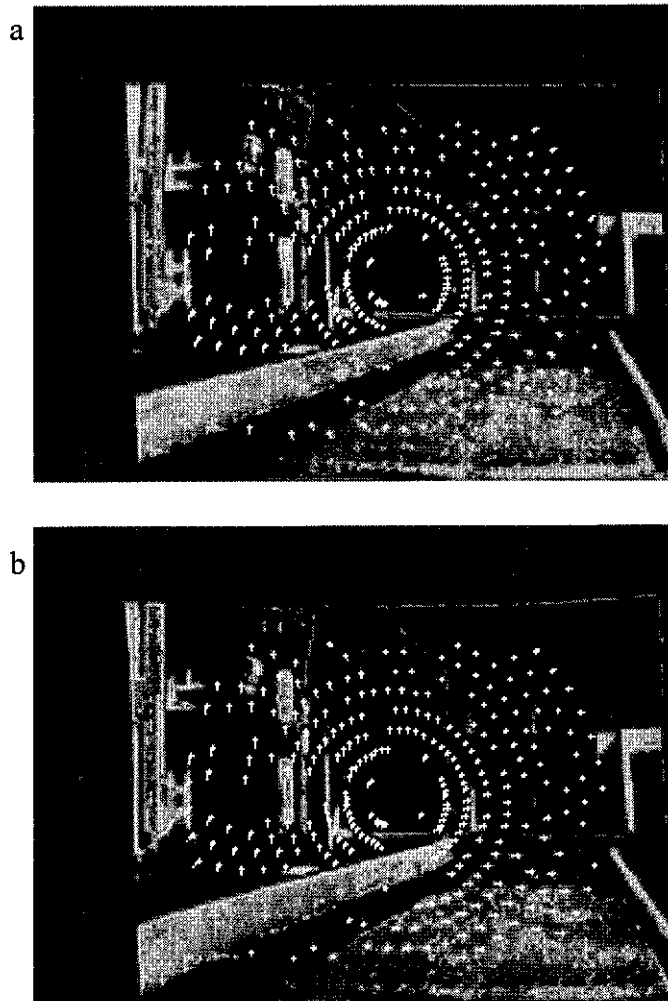
**Figura 7.21.** Componentes de Gabor de alta frecuencia,  $G_j$ . (a) Región de interés. (b) Imagen visible. (c) IR registrada con el modelo afín. (d) IR registrada con el modelo de tercer orden. (e) IR de referencia. (f), (g) y (h) Componentes de Gabor para las fusiones de las imágenes IR (c), (d) y (e) con la imagen visible.



**Figura 7.22.** Mapa de incrementos de energía generado para la zona de interés (7.21a), a partir de la imagen IR de referencia (7.20g) y la imagen visible (7.20b). El punto de mínima energía coincide con el de traslación nula (0, 0).



**Figura 7.23.** Curva de incrementos de energía  $\Delta E(r)$  en función de la distancia al punto mínimo, obtenida a partir del mapa de incrementos de la figura 7.22.



**Figura 7.24.** Campos vectoriales de desplazamientos locales, con respecto a la imagen de referencia óptima (7.20g), para las dos imágenes infrarrojas registradas de forma automática. (a) Imagen IR registrada mediante el modelo afín. (b) Imagen IR registrada mediante el modelo de tercer orden.



## 8. Aplicaciones del registro automático

En este capítulo se recogen algunos ejemplos de aplicación del método desarrollado para el registro automático de imágenes de bandas espectrales alejadas. En la sección 8.1 se muestran resultados del registro de imágenes visibles e infrarrojas, de especial importancia por constituir el principal objetivo de este trabajo. El caso particular del registro con imágenes SAR, de gran interés en aplicaciones de teledetección y que requiere el desarrollo de un procedimiento específico, se estudia en la sección 8.2.

Finalmente, en la sección 8.3, se citan otras posibles aplicaciones de los métodos de registro desarrollados.

### 8.1. Registro visible-IR

El principal objetivo del trabajo realizado era el desarrollo de un método automático para el registro de imágenes visibles e infrarrojas. A continuación se presentan dos ejemplos de la aplicación del método basado en calidad de casamiento con corrección de tercer orden para este tipo de registro (el primero aplicado a imágenes de entorno urbano, y el segundo a imágenes aéreas), así como las medidas de calidad de sus resultados.

En la figura 8.1 se muestra un par de imágenes IR-visible. La imagen infrarroja ha sido adquirida mediante una cámara térmica (8-12  $\mu\text{m}$ ) y la visible mediante una cámara de vídeo corriente. El resultado del registro automático y la fusión de la imagen IR con la visible se muestran en las figuras 8.1c y 8.1d, respectivamente.

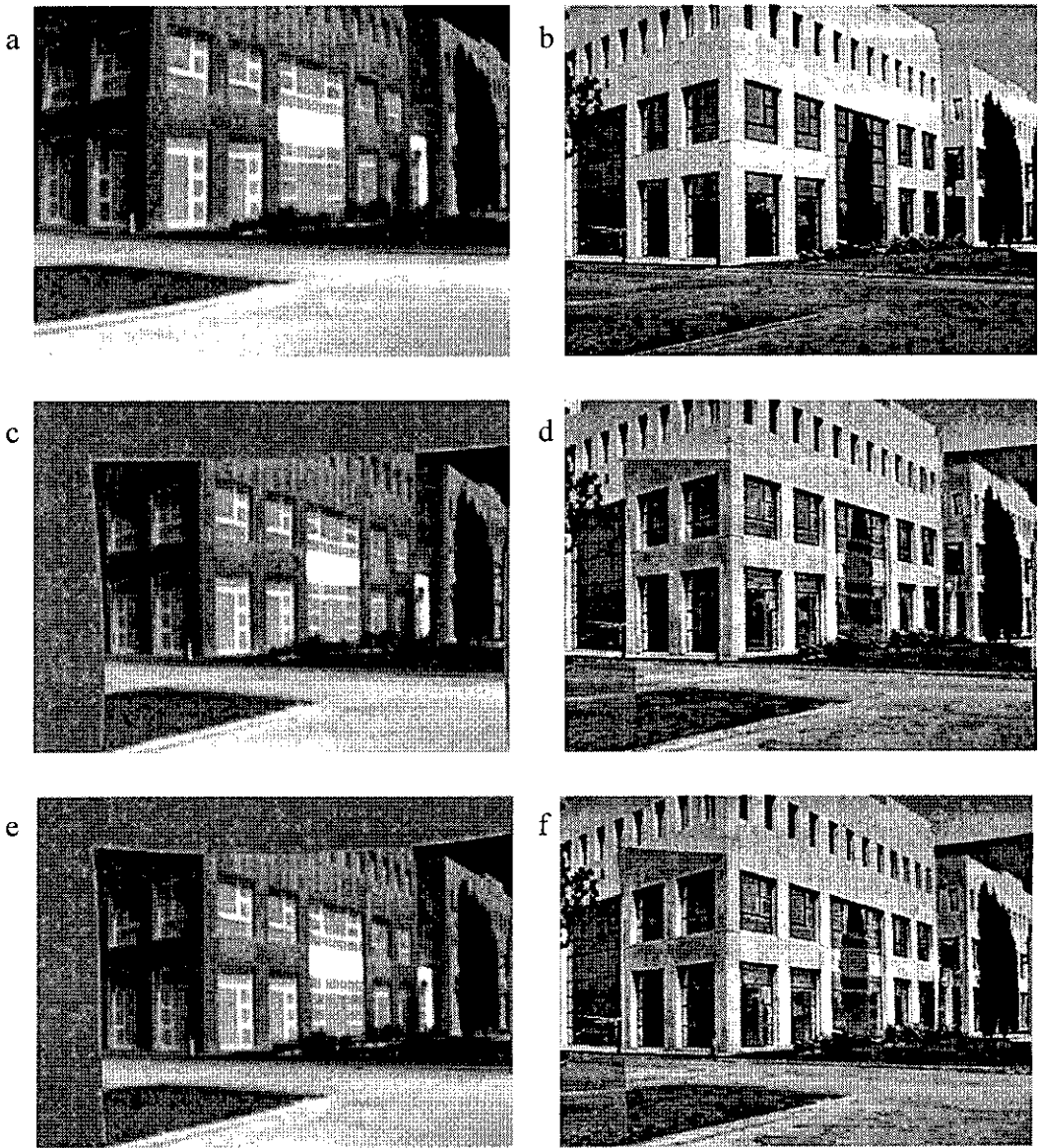
Para evaluar el resultado se ha realizado, además, un registro manual de las imágenes (figura 8.1e). La aplicación del criterio de calidad basado en Gabor genera el mapa y la curva de incrementos de energía recogidos en la figura 8.2.

El campo vectorial, detectado por el método de evaluación basado en el cálculo de desplazamientos locales se muestra en la figura 8.3. Los desplazamientos equivalentes para la imagen IR registrada automáticamente, obtenidos por los dos métodos de evaluación, pueden observarse en la tabla 8.1.

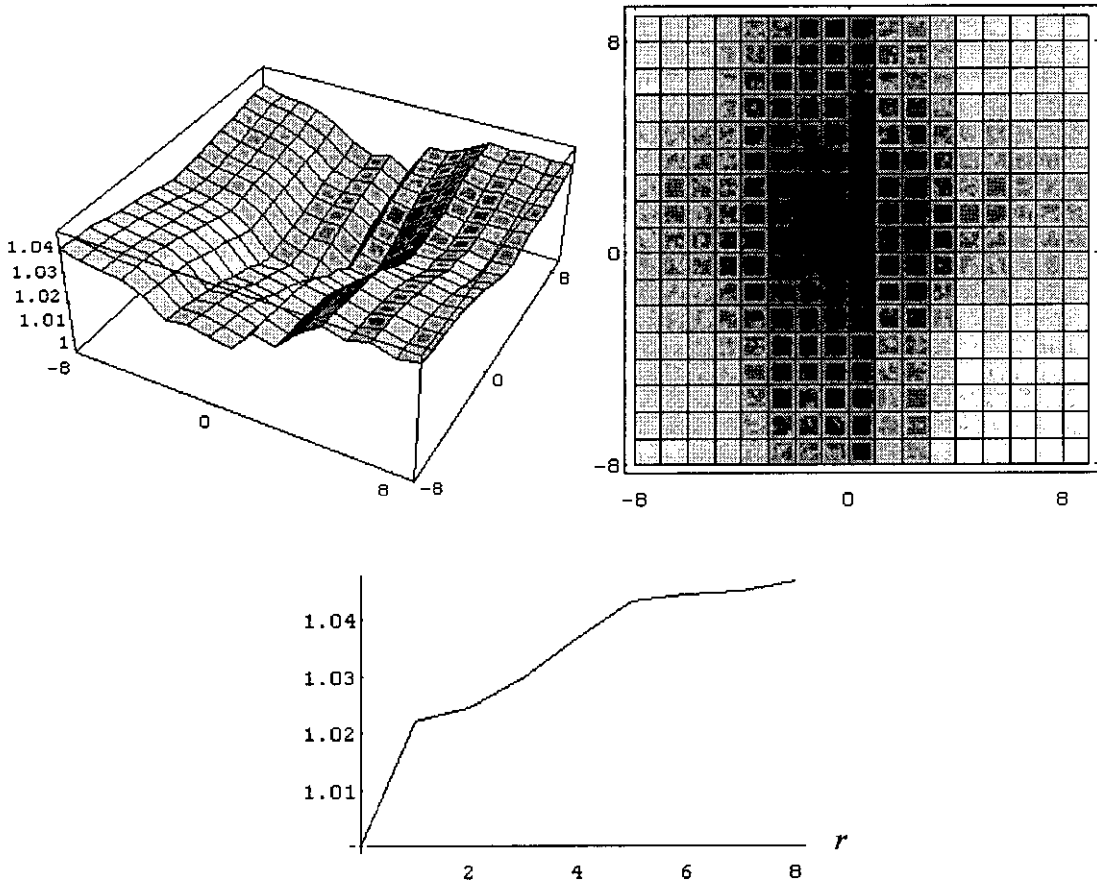


$\bar{r}_T$	0.084
$\ V\ $	0.625
$\bar{r}(V)$	0.530

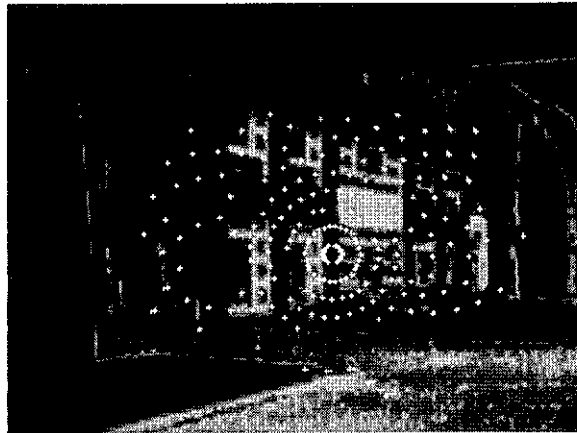
**Tabla 8.1.** Desplazamiento medio equivalente mediante el criterio de Gabor, media del módulo de los vectores del campo de desplazamientos locales y desplazamiento equivalente en función del campo vectorial, para el registro de las imágenes de la figura 8.1. Medidas en píxeles.



**Figura 8.1.** Dos imágenes, IR (a) y visible (b) de una misma escena. (c) Registro automático de (a) con (b) basado en calidad de casamientos con refinamiento de tercer orden, y (d) resultado de la fusión de (c) y (b). (e) Registro manual y (f) resultado de su fusión con (b).

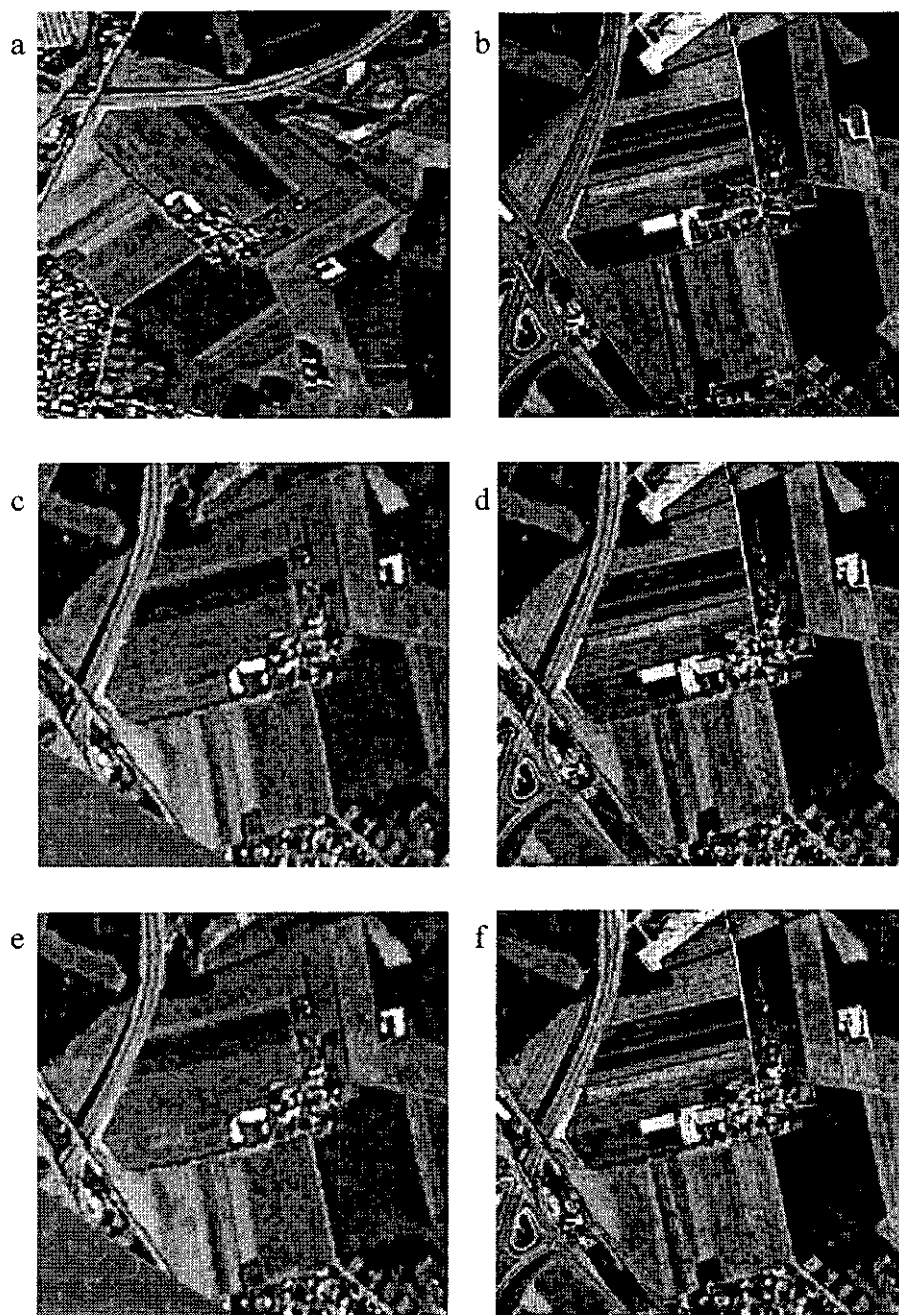


**Figura 8.2.** Mapa y curva de incrementos de energía para la evaluación del registro del par IR-visible de la imagen 8.1.



**Figura 8.3.** Campo vectorial de desplazamientos locales para la imagen 8.1c, registrada automáticamente, respecto de la registrada manualmente, 8.1e.

En la figura 8.4 se presentan los resultados del registro automático de un par de imágenes aéreas IR-visible.

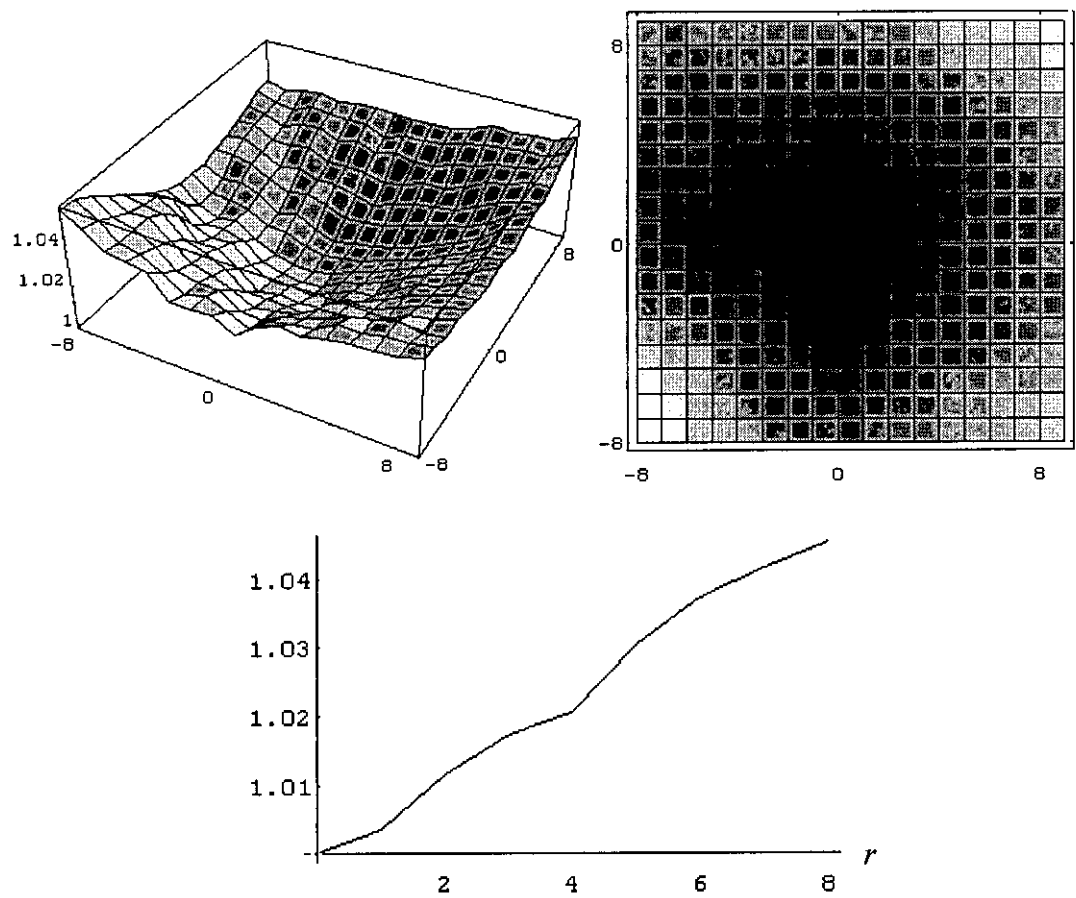


**Figura 8.4.** Registro de dos imágenes aéreas en bandas IR (a) y visible (b). (c) Registro automático de tercer orden y (d) fusión resultante. (e) Registro manual y (f) fusión.

En este caso, la aplicación del criterio de Gabor indica que el mínimo de energía se obtiene en el punto  $(-1, 0)$  del mapa de incrementos, mostrado en la figura 8.5. En consecuencia, la imagen registrada manual tendrá un desplazamiento equivalente no nulo. Como se recoge en la tabla 8.2, la diferencia entre el desplazamiento de la imagen registrada de forma automática con la manual es de aproximadamente 0.2 píxeles.

	$\bar{r}_T$ (píxeles)
manual	0.74
automática	0.92

**Tabla 8.2.** Desplazamientos equivalente para los registros manual y automático de la imagen IR 8.4a con la visible 8.4b, en función del criterio de Gabor.

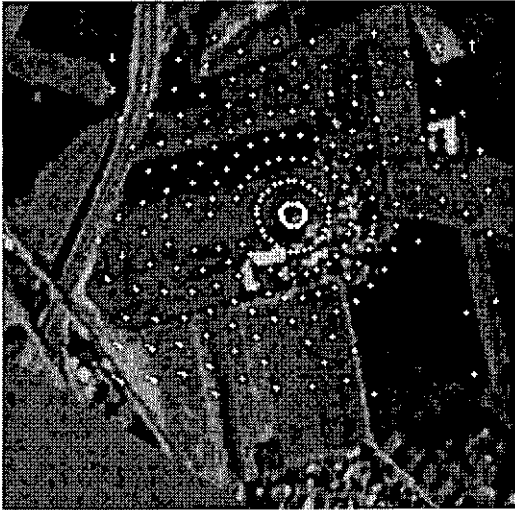


**Figura 8.5.** Mapa y curva de incrementos de energía para la evaluación del registro del par IR-visible de la figura 8.4.

Para la aplicación del método de evaluación basado en desplazamientos locales se ha tomado la imagen óptima, esto es, la manual después de aplicarle una traslación de  $-1$  píxel en dirección horizontal. El campo resultante se muestra en la figura 8.6 y el desplazamiento equivalente de la imagen registrada de forma automática está en la tabla 8.3.

$\ V\ $	1.00
$\bar{r}(V)$	0.82

**Tabla 8.3.** Desplazamientos equivalentes en función del campo vectorial de desplazamientos locales para la imagen 8.4c y la imagen registrada óptima, correspondiente a la traslación de (-1,0) píxeles de la imagen 8.4e. Medidas en píxeles.



**Figura 8.6.** Campo vectorial de desplazamientos locales para la imagen 8.4c, registrada automáticamente, respecto de la registrada óptima, correspondiente a la traslación de (-1, 0) píxeles de la registrada manualmente, 8.4e.

## 8.2. Registro visible/IR-SAR

### 8.2.1. Introducción

Las imágenes de Radar de Apertura Sintética (SAR) [2, 69, 70] son muy utilizadas en teledetección y control de recursos naturales, por su capacidad para diferenciar distintos tipos de terreno o niveles de humedad y por su operatividad en presencia de nubes. En ese sentido, es de gran interés combinar la información proporcionada por las imágenes SAR con la de imágenes visibles o infrarrojas, para lo cual será necesario registrarlas.

Las imágenes SAR, sin embargo, presentan menor resolución espacial y mayor nivel de ruido que las imágenes obtenidas por métodos ópticos en otras bandas espectrales (visible, IR), lo que dificulta el proceso de extracción de bordes e impide que puedan aplicarse directamente los métodos de registro basados en casamiento de segmentos. Aunque se han desarrollado múltiples métodos [71-74] para optimizar la extracción de bordes en las imágenes SAR, la calidad y el número de bordes detectados a resolución completa siguen siendo insuficientes para la aplicación de los métodos descritos en los capítulos anteriores, y es necesario desarrollar técnicas específicas para su registro.

Por otra parte, las imágenes utilizadas en el ámbito de la teledetección suelen ser de gran tamaño, por lo que su registro a resolución completa resulta muy costoso. Además, normalmente, las regiones de interés son pequeñas en relación al tamaño total de las imágenes, y es suficiente con que el registro sea preciso en esas pequeñas regiones. Por todo ello, hemos ampliado nuestros métodos de registro con el desarrollo de un procedimiento que permite registrar localmente imágenes visibles o IR con SAR de forma precisa, pero sin tener que realizar su registro global a resolución completa.

### 8.2.2. Descripción general del método de registro visible/IR-SAR

Los dos problemas principales que deben considerarse son el elevado nivel de ruido en la imagen SAR y el gran tamaño de las imágenes en comparación con la zona de interés. Para reducir la influencia de ambos factores, realizamos inicialmente un submuestreo de las imágenes. Se consigue, de este modo, reducir el nivel de ruido y el tamaño de las imágenes resultantes, que pueden ser entonces registradas de forma automática mediante un método basado en segmentos, o bien de forma manual. Finalmente, refinamos la transformación de

registro obtenida para las imágenes submuestreadas, de manera que sea precisa a resolución completa, para la región de interés elegida.

La función de registro para las imágenes submuestreadas,  $F$ , puede emplearse como aproximación a la función de registro a resolución completa. En efecto, dado un punto  $p_i = (x_i, y_i)$  de la imagen fuente (SAR), su punto correspondiente  $p'_i = (x'_i, y'_i)$  en la imagen destino (visible/IR) vendrá dado por:

$$p'_i \cong S \cdot F(S^{-1} \cdot p_i) \rightarrow \begin{cases} x'_i \cong s_x \cdot F_x\left(\frac{x_i}{s_x}, \frac{y_i}{s_y}\right) \\ y'_i \cong s_y \cdot F_y\left(\frac{x_i}{s_x}, \frac{y_i}{s_y}\right) \end{cases} \quad (8.1)$$

donde  $s_x$  y  $s_y$  indican los factores de submuestreo, y la matriz:

$$S^{-1} = \begin{pmatrix} 1/s_x & 0 \\ 0 & 1/s_y \end{pmatrix} \quad (8.2)$$

es la transformación de submuestreo utilizada, que debe ser incluida para compensar las diferencias de escala existentes entre los espacios métricos de las imágenes originales (a resolución completa) y las submuestreadas. El error cometido por la utilización de  $F$ , en lugar de la función de registro real, es proporcional a las constantes de submuestreo, tal y como puede verse en las fórmulas de (8.1).

Para refinar  $F$  debe encontrarse una corrección,  $\Delta F$ , que compense las imprecisiones de  $F$  a resolución completa. La transformación entre los espacios fuente y destino tendrá entonces la forma:

$$p'_i = S \cdot F(S^{-1} \cdot p_i) + \Delta F(p_i) \quad (8.3)$$

Para calcular  $\Delta F$  no puede utilizarse la técnica de refinado descrita en el capítulo 6, puesto que no es posible realizar una extracción de segmentos a resolución completa en la imagen SAR, y deben buscarse procedimientos alternativos. Por ello, hemos desarrollado un método de corrección, basado en acumulación, que utiliza un criterio estadístico para la determinación de la mejor transformación de registro local.

### 8.2.3. Refinamiento del registro mediante acumulación

Para determinar la mejor corrección a la función de registro aproximada, transformamos sobre la imagen SAR todos los segmentos detectados en la región de interés de la imagen visible/IR, y obtenemos la transformación local que mejor los ajuste a la región correspondiente de la imagen SAR.

Como no puede realizarse una detección de bordes, no puede buscarse un segmento correspondiente para cada segmento transformado. Sin embargo, es de esperar que cada segmento transformado se encuentre en las proximidades de un borde de la imagen SAR, dentro de los límites impuestos por el error cometido al emplear las funciones de registro aproximadas. Si esos límites son pequeños, se puede realizar un barrido sobre los parámetros de registro para encontrar la corrección más probable para cada segmento. La acumulación sobre el espacio de parámetros de las probabilidades obtenidas para todos los segmentos de la zona, permite determinar la corrección global más adecuada para toda la región de interés. Deben especificarse, por tanto, un modelo para  $\Delta F$  y una función de peso adecuada.

Para determinar el modelo de función de corrección,  $\Delta F$ , debe estudiarse qué tipo de imprecisiones causa la aplicación de la función de registro aproximada  $F$  en la transformación de los segmentos. Podemos usar para ello las fórmulas empleadas en el capítulo 4 para obtener las precisiones asociadas a cada uno de los parámetros de registro afines. Estas fórmulas permiten obtener los valores mínimos que pueden calcularse para cada uno de los parámetros afines en función del error en la localización de los puntos (asociado allí a la discretización del espacio imagen, y aquí a la imprecisión en la posición debida al submuestreo). Como veremos a continuación, mediante el análisis de esos valores mínimos, únicamente es necesario realizar una corrección en las coordenadas de traslación, ya que los demás parámetros quedan determinados con sobrada precisión, incluso si se calculan sobre las imágenes submuestreadas.

En efecto, es previsible que aparezcan errores de traslación, puesto que el submuestreo implica una incertidumbre en la posición proporcional a la constante de escalado que se utilice. De hecho, el resultado del cálculo de la transformación de registro aproximada sobre las imágenes submuestreadas es equivalente al cálculo de la misma sobre las imágenes a resolución completa realizando una cuantización del espacio en unidades de longitud  $dl'$ , iguales a la incertidumbre en la localización causada por el submuestreo, esto es:



$$\begin{aligned}
dl' &\approx dl'_x, dl'_y \\
dl'_x &= s_x \cdot dl = s_x \text{ píxeles} \\
dl'_y &= s_y \cdot dl = s_y \text{ píxeles}
\end{aligned} \tag{8.4}$$

donde  $dl = 1$  píxel es la resolución de la imagen y  $s_x, s_y$  son los factores de submuestreo utilizados.

En el caso de la traslación, el error a resolución completa se corresponde directamente con las incertidumbres de localización (véase la fórmula 4.6):

$$\begin{aligned}
dt_x &= dl'_x = s_x \text{ píxeles} \\
dt_y &= dl'_y = s_y \text{ píxeles}
\end{aligned} \tag{8.5}$$

Las incertidumbres para los demás parámetros se calculan aplicando las fórmulas 4.7, 4.8 y 4.9:

$$\begin{aligned}
d\alpha &= \tan^{-1} \left( \sqrt{\frac{L^2}{(dl')^2} - 1} \right) - \tan^{-1} \left( -1 + \sqrt{\frac{L^2}{(dl')^2} - 1} \right) \approx \\
&\approx \tan^{-1} \left( \sqrt{\frac{L^2}{s^2} - 1} \right) - \tan^{-1} \left( -1 + \sqrt{\frac{L^2}{s^2} - 1} \right)
\end{aligned} \tag{8.6}$$

$$ds = \sqrt{1 + \left( \frac{dl'}{L} \right)^2} - 1 \approx \sqrt{1 + \left( \frac{s}{L} \right)^2} - 1 \tag{8.7}$$

$$ds_k = \frac{(dl')^2}{L^2 + (dl')^2 - dl' \cdot \sqrt{L^2 + (dl')^2}} \approx \frac{s^2}{L^2 + s^2 - s \cdot \sqrt{L^2 + s^2}} \tag{8.9}$$

donde  $s \approx s_x, s_y$ .

Para las imágenes de prueba utilizadas (imágenes aéreas y de satélite, en bandas visible, IR y SAR) el tamaño es del orden de  $10^4 \times 10^4$  píxeles, y han sido submuestreadas a  $512 \times 512$  píxeles para su prerregistro, realizado de forma manual. Por tanto, suponiendo para  $L$  y  $s$  los valores:

$$\begin{aligned}
L &\approx 10^4 \text{ píxeles} \\
s &= s_x = s_y \approx 10
\end{aligned} \tag{8.10}$$

se obtienen para los parámetros de la transformación de registro aproximada  $F$  en la imagen a resolución completa las incertidumbres siguientes:

$$\begin{aligned} dt_x &= dt_y = 10 \text{ píxeles} \\ ds_x &= ds_y = 5 \cdot 10^{-7} \\ d\alpha &= 10^{-6} \text{ rad} = 6 \cdot 10^{-5}^\circ \\ ds_k &= 10^{-6} \end{aligned} \tag{8.11}$$

Como puede verse, sólo la incertidumbre en traslación es relevante, ya que los demás valores son demasiado pequeños como para ser apreciables.

La consecuencia principal de estos resultados es que, si la deformación en las imágenes submuestreadas ha sido restaurada correctamente por el modelo, entonces es suficiente con corregir las ventanas a resolución completa en traslación. Esto simplifica sobremanera la forma de la función  $\Delta F$  y el modo de obtenerla, ya que únicamente será necesario desplazar los segmentos alrededor de la posición determinada por  $F$  para encontrar su localización más probable.

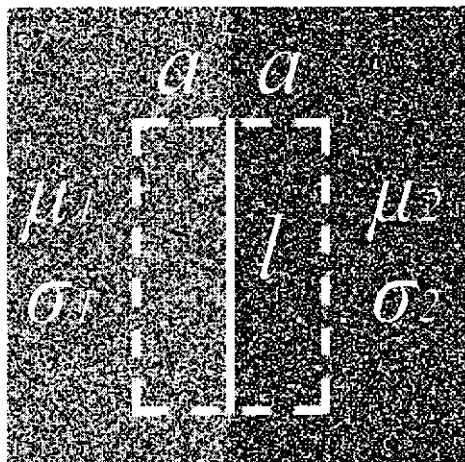
Como solamente es necesario calcular una función global de dos coordenadas, y su cálculo se fundamenta en evaluaciones locales de la posición de los segmentos, la aplicación de técnicas basadas en acumulación resulta muy adecuada. Sólo resta, por tanto, definir una función de peso adecuada para determinar la localización más probable de cada segmento.

### 8.2.3. Función de peso para la localización de segmentos

Puesto que no es posible extraer bordes ni detectar segmentos con calidad suficiente para el registro en la imagen SAR, debe buscarse otra forma de estimar la posición más adecuada para cada segmento transformado sobre ella. Nuestro método se basa en cuantificar la diferencia existente entre las regiones que separa el segmento. En efecto, como cada borde de la imagen se corresponde, en el fondo, con la frontera existente entre dos regiones de diferentes características, puede utilizarse la diferencia existente entre las regiones que separa el segmento como medida de la probabilidad de localizar un borde en la posición en la que se encuentra.

Para modelizar las regiones separadas por el segmento, supondremos que presentan una distribución gaussiana de niveles de gris. La estimación de los niveles de gris medios ( $\mu_l$  y

$\mu_2$ ) y sus varianzas ( $\sigma_1$  y  $\sigma_2$ ) se calculan analizando dos áreas rectangulares (a las que llamamos *regiones de medida*) situadas a cada uno de los lados del segmento, como se muestra en la figura 8.7. La posición más adecuada para el segmento será aquella en la que la discriminación entre las poblaciones de las regiones de medida sea máxima. En cualquier otra posición se estarán incluyendo en una de las poblaciones píxeles de la otra, lo que reducirá las diferencias entre los parámetros que las caracterizan.



**Figura 8.7.** Regiones de medida para refinar la posición de un segmento, de longitud  $l$ , sobre la imagen SAR. Las regiones de medida son de forma rectangular, de altura  $l$  y anchura  $a$ , y las poblaciones de niveles de gris contenidas en ellas se modelan mediante distribuciones gaussianas,  $(\mu_1, \sigma_1)$  y  $(\mu_2, \sigma_2)$ .

Es necesario, por tanto, dividir los puntos de las regiones de medida en dos grupos (o *clusters*), de forma que la discriminación entre sus distribuciones de población sea óptima. La discriminación será máxima cuando el segmento se encuentre justo en la frontera entre las dos zonas de distribuciones de gris diferentes de la imagen SAR, de forma que en cada una de las regiones de medida sólo entren puntos de una de las dos zonas.

Para calcular el nivel de diferenciación de los grupos hemos empleado la función discriminante de Fisher [10], que se utiliza para determinar la dirección de un vector que divida de forma óptima un conjunto de vectores de atributos (*feature vectors*),  $\{\mathbf{x}_j, j = 1, \dots, N\}$ , en dos grupos,  $C_1$  y  $C_2$ . El criterio de optimización se basa en la maximización del cociente:

$$D(\mathbf{v}) = \frac{\mathbf{v} \mathbf{S}_b \mathbf{v}^T}{\mathbf{v} \mathbf{S}_w \mathbf{v}^T} \quad (8.12)$$

donde  $S_b$  y  $S_w$  son las matrices de covarianza íter- e intra-población, y  $v$  es un vector que designa una dirección sobre la que se proyecta ortogonalmente el espacio de atributos.

En general, dado un conjunto de vectores de atributos  $\{x_j, j = 1, \dots, N\}$  clasificados en  $K$  grupos,  $C_k$ , la matriz de covarianza intra-población,  $S_w$ , se define como:

$$S_w = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k} \sum_{x_i \in C_k} (x_i - \mu_k)(x_i - \mu_k)^T \quad (8.13)$$

donde  $N_k$  y  $\mu_k$  son el número de elementos y el vector de atributos medio del grupo  $C_k$ , respectivamente.

La matriz de covarianza íter-población,  $S_b$ , se define de la siguiente forma:

$$S_b = \frac{1}{K} \sum_{k=1}^K (\mu_k - \mu_0)(\mu_k - \mu_0)^T \quad (8.14)$$

donde  $\mu_0$  es el vector medio del conjunto total de vectores de atributos:

$$\mu_0 = \frac{1}{N} \sum_{i=1}^N x_i \quad (8.15)$$

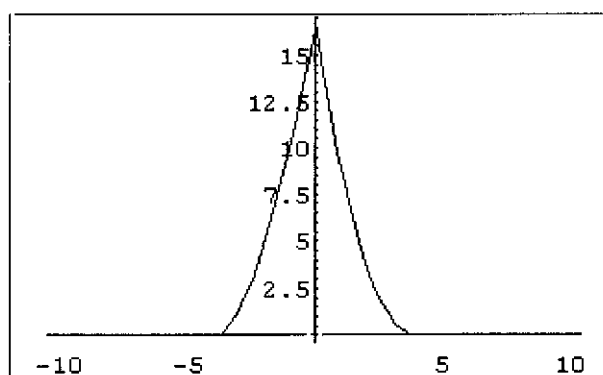
En nuestro caso, los vectores de atributos tienen dimensión uno (el nivel de gris de los puntos de las regiones de medida a cada uno de los lados del segmento), y el número de grupos es dos. En tales condiciones la función de peso (8.14) tiene la forma siguiente:

$$D = \frac{(\mu_1 - \mu_0)^2 + (\mu_2 - \mu_0)^2}{\sigma_1^2 + \sigma_2^2} \quad (8.16)$$

donde  $\mu_1$ ,  $\mu_2$  y  $\sigma_1^2$ ,  $\sigma_2^2$  indican los valores medios y las varianzas de las distribuciones de nivel de gris de las regiones de la imagen SAR situadas a cada uno de los lados del segmento, y donde  $\mu_0$  vale:

$$\mu_0 = \frac{\mu_1 + \mu_2}{2} \quad (8.17)$$

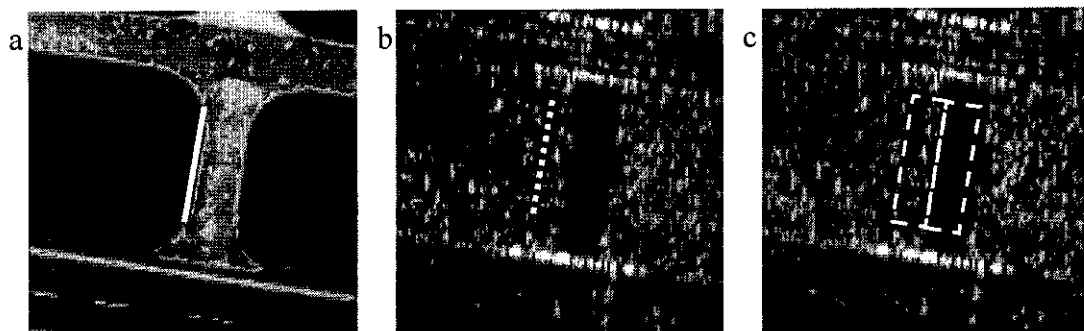
La forma de la función discriminante  $D$  se muestra en la figura 8.8, para un conjunto de valores arbitrario.



**Figura 8.8.** Variación del valor del discriminante de Fisher,  $D$ , según se modifica la posición del segmento perpendicularmente a la frontera de separación de dos regiones de medias y varianzas diferentes. El máximo se produce cuando la posición del segmento coincide con la frontera de separación de las regiones. En este ejemplo  $\mu_1 = \sigma_1^2 = 100$ ,  $\mu_2 = \sigma_2^2 = 200$ ,  $a = 4$  píxeles y  $l = 12$  píxeles.

La anchura,  $a$ , de las regiones de medida es un parámetro variable en nuestra implementación de este método de registro. Su valor está relacionado con el tamaño de los mínimos objetos de interés dentro de las imágenes a registrar. Por una parte, debe utilizarse una anchura para las regiones de medida lo suficientemente grande como para cubrir una zona en la que los valores de la media y la varianza puedan estimarse correctamente; por otra, la anchura debe ser lo suficientemente pequeña como para que las regiones de medida queden contenidas dentro de zonas uniformes. Para las imágenes de prueba en las que hemos utilizado este método de registro (principalmente imágenes aéreas de aeropuertos de unos 2.5 m/píxel de resolución), se ha comprobado experimentalmente que el valor más adecuado está alrededor de los tres o cuatro píxeles. Valores más pequeños no permiten una evaluación adecuada de los parámetros de las distribuciones en las regiones de medida, y valores mayores incluyen en las regiones de medida elementos de regiones diferentes a las inmediatamente adyacentes al segmento, con lo que el error de los valores estimados para los parámetros de las distribuciones aumenta.

En la figura 8.9 se muestra el proceso de ajuste de un segmento detectado en una imagen visible sobre la imagen SAR correspondiente. La posición inicial del segmento (figura 8.9b) se determina mediante la transformación de registro aproximada,  $F$ . Después se busca localmente la traslación que maximiza el valor de la función  $D$ , que determina la posición más probable para el segmento (figura 8.9c).



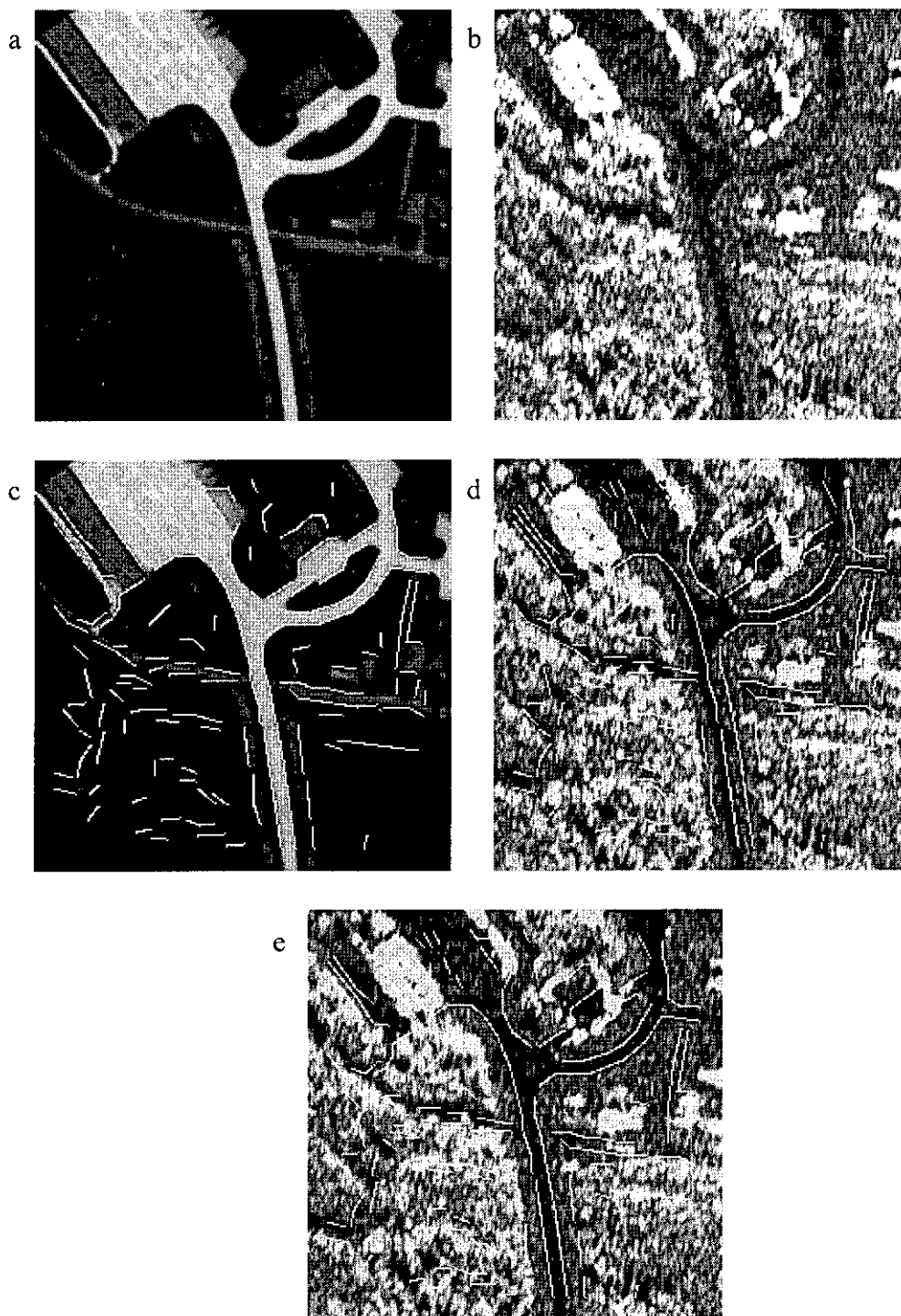
**Figura 8.9.** Ajuste fino sobre la imagen SAR de un segmento transformado mediante la función de registro aproximada. (a) Segmento detectado en la imagen visible. (b) Segmento transformado sobre la imagen SAR con la función de transformación aproximada. (c) Las dos regiones de medida a los lados del segmento transformado, en la posición en la que el nivel de discriminación entre sus poblaciones es máximo.

Para corregir globalmente la región de interés, se acumulan en una matriz bidimensional las probabilidades asociadas a cada una de las traslaciones posibles para cada uno de los segmentos de la región (detectados en la imagen visible). El máximo de acumulación se corresponderá con el ajuste en traslación más adecuado para la región elegida.

Como límites máximos para las coordenadas de traslación, alrededor de la posición inicial, tomamos el doble del error asociado al submuestreo. Puesto que únicamente es necesario corregir dos parámetros de la transformación, y el número de segmentos en la región de interés no suele ser muy elevado (típicamente del orden de 100), la técnica de acumulación, en este caso, no requiere mucha potencia de proceso.

En la figura 8.10 se muestra un ejemplo de aplicación del método de registro visible/IR-SAR desarrollado.

Como puede verse en la figura 8.10e, las posiciones correctas de muchos de los segmentos visibles sobre la imagen SAR no se corresponden exactamente con las relaciones que establecería un operario humano al registrar las imágenes manualmente. Este hecho se pone de manifiesto al aplicar el criterio de Gabor para la evaluación del registro (tabla 8.4), que, en este caso, asigna un incremento de energía menor al registro automático que al manual, lo que demuestra el correcto funcionamiento del método desarrollado.

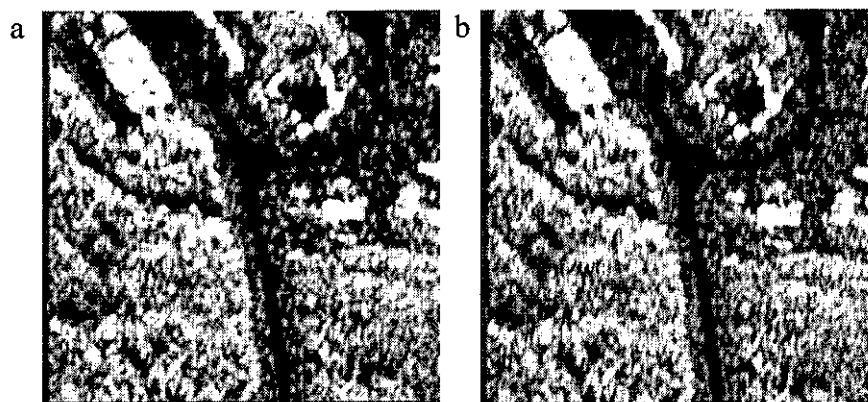


**Figura 8.10.** Registro visible-SAR mediante el método propuesto. (a) Región de interés en la imagen visible. (b) Región correspondiente de la imagen SAR, determinada mediante la transformación de registro aproximada. (c) Segmentos detectados en la imagen visible. (d) Segmentos transformados sobre la imagen SAR utilizando la transformación aproximada. (e) Región SAR re-extraída de la imagen completa utilizando la transformación corregida. Como puede verse, los segmentos transformados coinciden ahora de forma precisa con los bordes de la imagen SAR.

	$\bar{r}_T$ (píxeles)
manual	6.99
automática	0.44

**Tabla 8.4.** Desplazamientos equivalentes para los registros manual y automático de la imagen SAR 8.10a con la visible 8.10b, en función del criterio de Gabor.

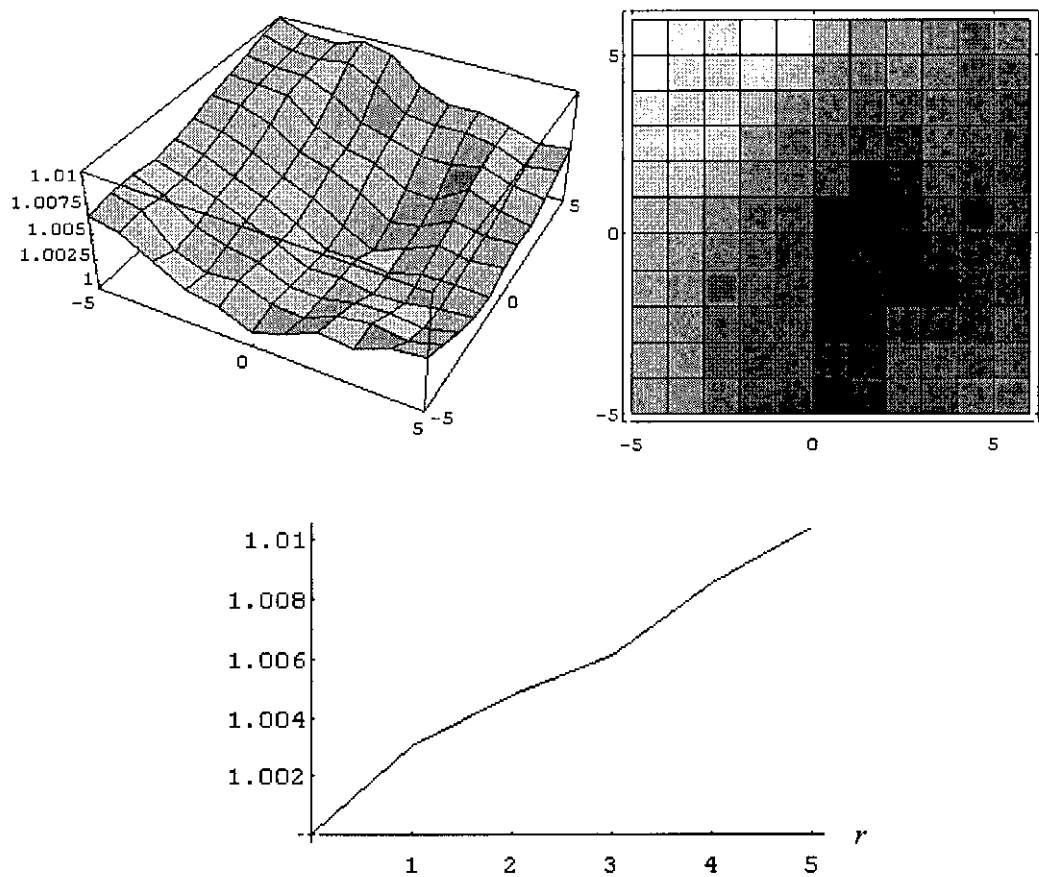
En la figura 8.11 pueden compararse los resultados de los registros automático y manual de las imágenes 8.10a y 8.10b. El valor de energía mínimo se obtiene para un desplazamiento de (1, 0) píxeles de la imagen registrada automáticamente (figura 8.12). La curva de incrementos obtenida crece de forma muy lenta, lo que significa que los desregistros pequeños entre las imágenes 8.10a y 8.10b son difíciles de apreciar visualmente (compárense los resultados con los demás ejemplos mostrados en la memoria). Este lento crecimiento también explica el elevado valor de desregistro obtenido para la imagen registrada manual



**Figura 8.11.** Registro (a) automático y (b) manual de la imagen SAR 8.10a con la visible 8.10b, utilizados para la determinación de la calidad en función del criterio de Gabor.

El método de registro visible/IR-SAR descrito en esta sección ha sido utilizado dentro de un proyecto internacional para el Centro de Satélites de la Unión Europea Occidental (WEUSC), que incluye tareas de registro y fusión de imágenes de aeropuertos adquiridas en diferentes bandas espectrales. El procedimiento de registro está integrado en una aplicación que permite al usuario determinar, sobre la versión submuestreada de las imágenes prerregistradas, la zona que desea examinar a resolución completa. El programa utiliza entonces el método de refinamiento desarrollado para realizar las correcciones necesarias.





**Figura 8.12.** Mapa de incrementos de energía de la imagen SAR registrada automáticamente con la visible. Debajo, curva de desplazamientos equivalentes.

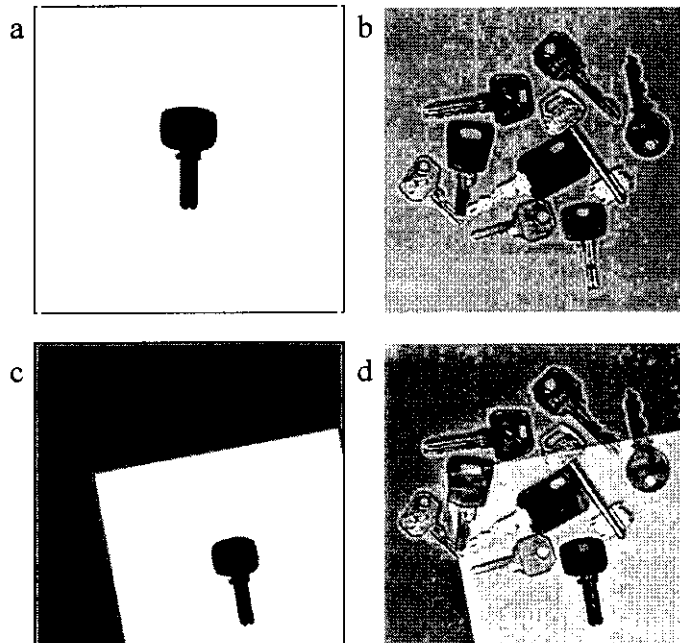
### 8.3. Otras aplicaciones

El método desarrollado para el registro de imágenes de bandas espectrales alejadas, debido a la generalidad con la que ha sido concebido, puede emplearse también para otras tareas que normalmente se realizan mediante métodos específicos. A continuación se citan algunas posibles aplicaciones.

#### 8.3.1. Reconocimiento y detección de formas

Tanto los métodos de registro basados en acumulación como los que utilizan calidad de casamiento de segmentos pueden emplearse para la detección de formas, utilizando como imagen fuente la forma que se desea reconocer y como destino la imagen donde debe ser buscada.

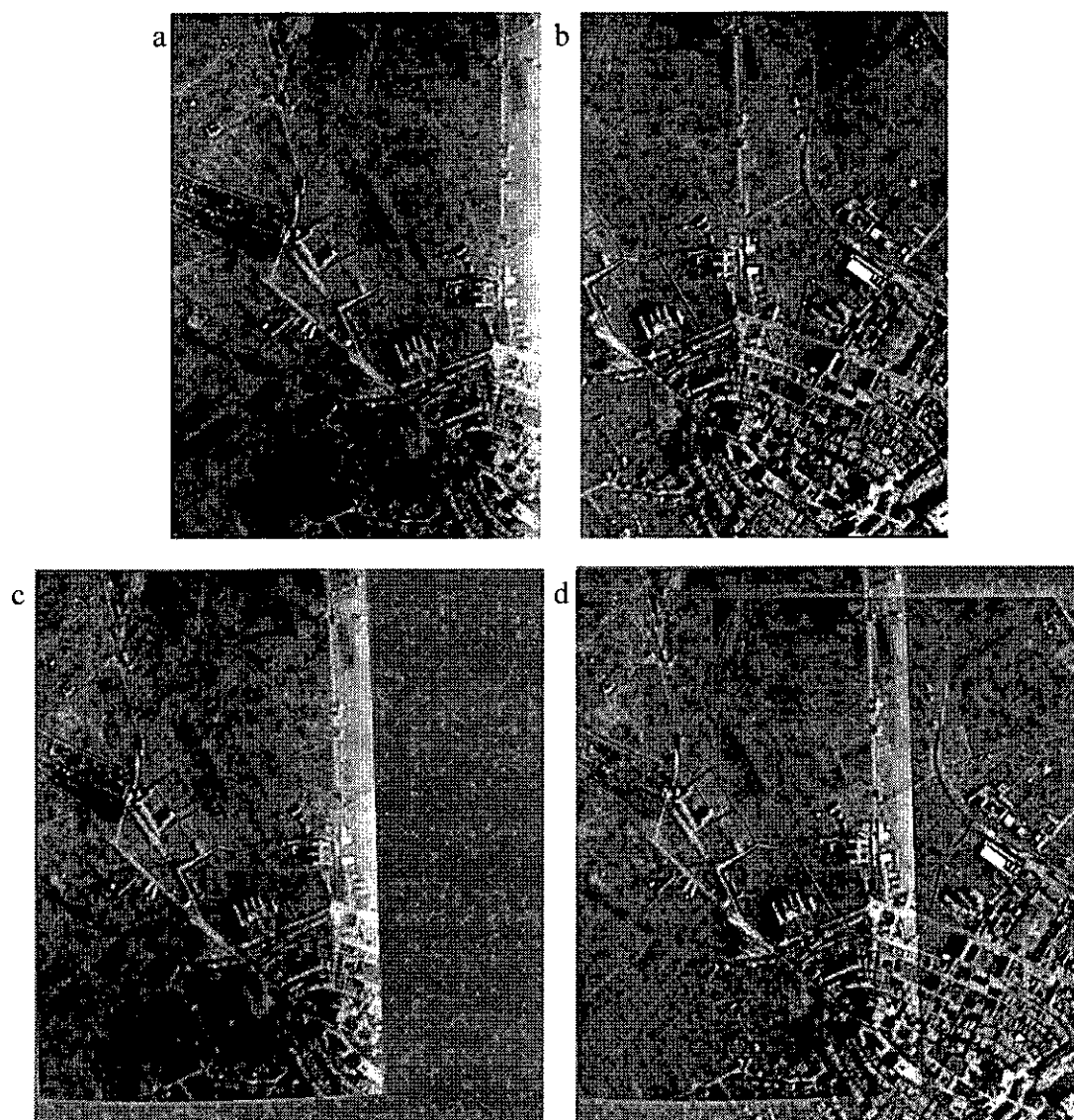
En la figura 8.13 se muestran la silueta de una llave y una imagen en la que debe ser identificada. Para este ejemplo se ha utilizado el registro afin basado en calidad de casamiento de segmentos, reduciendo severamente las tolerancias para la función de calidad, de forma que sólo los casamientos muy precisos obtuviesen valoraciones altas. La reducción de las tolerancias es imprescindible en este caso, dado el elevado grado de parecido que presentan las siluetas de las llaves de la imagen de búsqueda.



**Figura 8.13.** Aplicación del registro al reconocimiento de formas. (a) Silueta del objetivo y (b) imagen de búsqueda. (c) Silueta registrada y (d) su fusión con (b).

### 8.3.2. Mosaicos de imágenes

Nuestro método de registro para imágenes de diferentes bandas espectrales también puede emplearse para el registro de imágenes del mismo tipo. Es el caso del registro de imágenes para la generación de mosaicos.

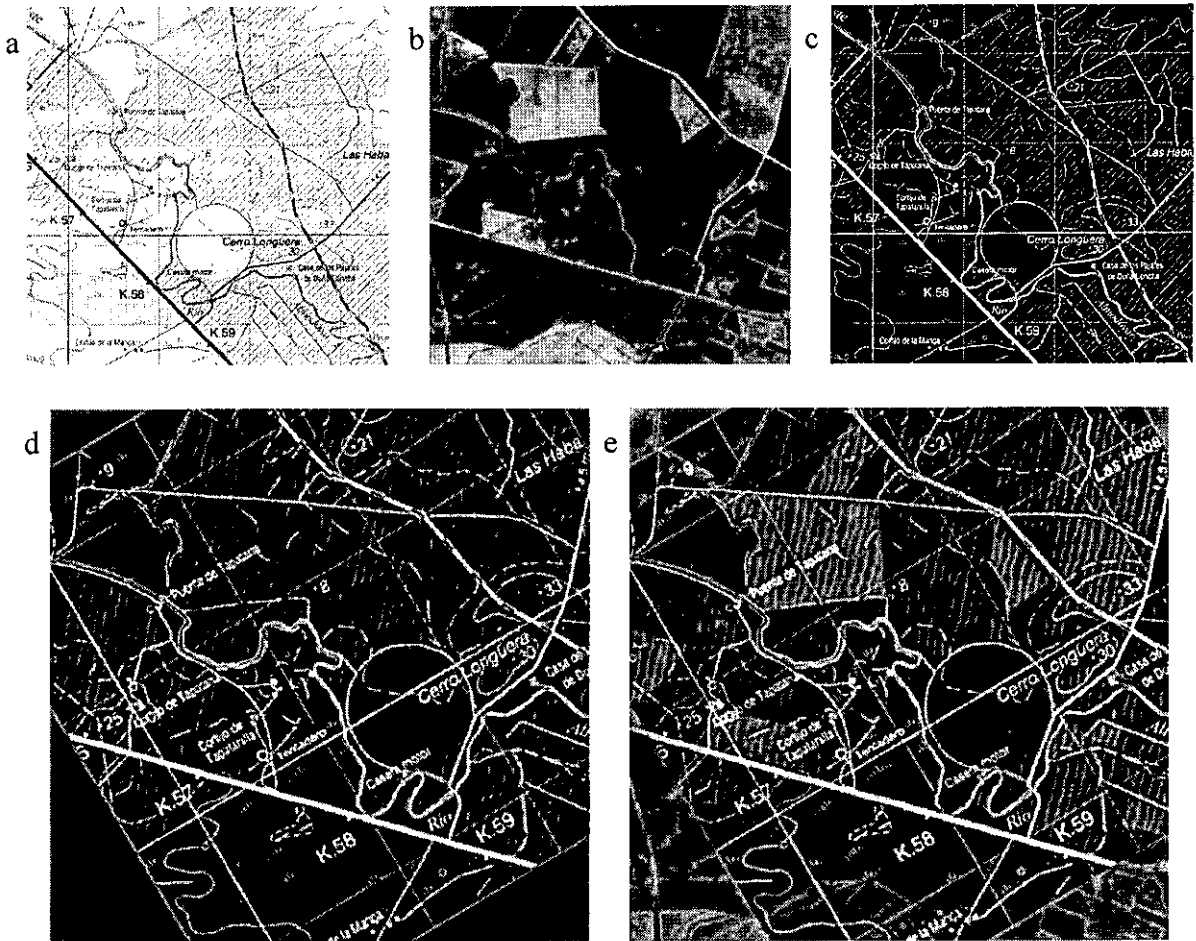


**Figura 8.14.** Aplicación del registro a la generación de mosaicos de imágenes. (a) y (b) Dos fotografías de la parte norte de Vitoria. (c) Registro de (a) con (b) y (d) fusión de (c) con (b).

En la figura 8.14 se muestran dos fotografías aéreas (a y b) proporcionadas por el Centro de Estudios Ambientales de Vitoria, que deben ser ensambladas por su zona común, y los resultados de su registro (c) y fusión (d) mediante un polinomio de tercer orden. Puede encontrarse más información sobre este trabajo en el artículo de Miravet, Santamaría, Coiras y Ureña [75].

### 8.3.3. Registro de mapas y fotografías

También puede utilizarse nuestro método para el registro de imágenes procedentes de sensores de naturaleza completamente diferente, y para el registro de imágenes de tipos muy distintos. Es el caso del registro de mapas con fotografías aéreas o de satélite, muy útil para aplicaciones cartográficas y de teledetección.



**Figura 8.15.** Aplicación del registro a la fusión de mapas y fotografías. (a) Mapa y (b) fotografía aérea de una misma zona. (c) Mapa umbralizado, utilizado como imagen de bordes. (d) Mapa registrado con la fotografía y (e) resultado de la fusión de ambos.

Esta técnica de registro puede emplearse en combinación con el método de extracción automática de redes lineales en mapas a color que hemos desarrollado [80]. Se dispone así de unas potentes herramientas para el análisis y el procesado conjunto de información cartográfica y telemétrica.



## 9. Conclusiones

El trabajo realizado, descrito en esta memoria, se ha orientado al desarrollo de procedimientos automáticos de registro para imágenes de bandas espectrales alejadas. A continuación se enumeran las principales conclusiones obtenidas:

1. La utilización de métodos de registro basados en casamiento de segmentos ajustados a los bordes constituye la opción más sólida para el registro de dos imágenes de una misma escena adquiridas por sensores de naturaleza diferente.
2. Mediante la asociación de estructuras formadas por los segmentos detectados puede determinarse un conjunto de potenciales transformaciones de registro. Para ello, hemos utilizado principalmente agrupaciones de segmentos en tríos (de modo que formen triángulos), que resultan especialmente adecuadas porque su emparejamiento define directamente una transformación afín completa. También se han utilizado parejas de segmentos y segmentos aislados en algunas aplicaciones particulares. Estas estructuras, más sencillas, si bien no generan directamente transformaciones completas, permiten reducir el tamaño de los conjuntos de elementos que deben casarse y, por tanto, la complejidad de los cálculos a realizar.
3. Para dar una significación y realizar un prefiltrado de las transformaciones obtenidas mediante el casamiento de estructuras formadas por segmentos, hemos utilizado un modelo de transformación afín cuyos parámetros se corresponden directamente con magnitudes geométricas fácilmente interpretables. Hemos demostrado que este modelo es completo, en el sentido de que es equivalente a la descomposición en valores singulares de las matrices que caracterizan a las transformaciones afines.
4. Para la selección de la transformación de registro óptima entre el conjunto de transformaciones candidatas, hemos seguido dos líneas principales:
  - La primera se basa en la idea de que la transformación global óptima debe corresponderse con la que optimice globalmente los casamientos individuales entre estructuras. Esta premisa constituye el fundamento sobre el que hemos desarrollado los métodos basados en acumulación, que, si bien no permiten resolver el problema del registro genérico, resultan de gran utilidad en varios casos particulares.

- La segunda supone que sólo puede determinarse la transformación global óptima mediante la evaluación del casamiento global resultante de la aplicación de las transformaciones de registro candidatas a los conjuntos de segmentos. Esta línea conduce al desarrollo de métodos basados en calidad de casamiento entre segmentos, mediante los cuales se ha conseguido resolver el problema del registro para cualquier tipo de imágenes.
5. Para evaluar el casamiento total resultante de la aplicación de una transformación de registro candidata a los conjuntos de segmentos, hemos empleado dos criterios diferentes para valorar el casamiento entre un segmento fuente y uno destino.
- El primero se basa en idea de que el casamiento será mejor cuanto mayor sea la superposición entre los segmentos. Hemos comprobado la validez de este criterio para el registro de imágenes similares, aunque no resulta útil para la resolución del problema de registro entre imágenes diferentes.
  - El segundo modo de evaluación supone que únicamente puede emplearse la pertenencia de los segmentos a una misma recta del plano imagen para su casamiento. Este criterio es consecuencia directa de suponer que si dos segmentos casan, es porque se corresponden con el mismo borde físico de la escena. Por tanto, deben pertenecer a una misma recta de la imagen, resultante de la proyección del borde físico de la escena sobre el plano de imagen. Mediante esta idea se ha definido una distancia entre segmentos que mide la posibilidad de pertenencia de un segmento a la recta definida por otro. A partir de la distancia, hemos definido una función de calidad de casamiento entre conjuntos de segmentos, que puede emplearse tanto para la evaluación de potenciales transformaciones de registro, como para encontrar correspondencias entre segmentos o detectar segmentos colineales.
6. Mediante la utilización de triángulos para la determinación de transformaciones candidatas, y el empleo de la función de calidad de casamiento para la valoración de aquellas, se ha construido un método robusto para el registro afin de imágenes de bandas espectrales separadas. El procedimiento desarrollado permite reconstruir deformaciones de magnitud arbitraria (esto es, no impone límites a los valores de sus parámetros) y puede aplicarse, en general, a imágenes de cualquier naturaleza.
7. Adicionalmente, hemos desarrollado un método de refinamiento de orden superior para el registro de imágenes que presenten deformaciones no lineales. Este método se

fundamenta en el ajuste iterativo del conjunto de segmentos fuente sobre el conjunto destino para corregir progresivamente los parámetros de la transformación de registro.

8. Se han aplicado los métodos de registro desarrollados a los casos prácticos más frecuentes, como el registro visible-IR o el registro visible/IR-SAR. Este último caso, de gran importancia, ha requerido un tratamiento especial, debido a la imposibilidad de realizar directamente una extracción de segmentos en la imagen SAR. Por ello, hemos desarrollado otro procedimiento que obtiene una transformación inicial de registro a partir de las imágenes submuestreadas, cuyos parámetros se refinan posteriormente en la región de interés utilizando acumulación.
9. Para la valoración de nuestros procedimientos de registro se han desarrollado dos métodos de evaluación diferentes.
  - El primer método, basado en el análisis del resultado de la fusión de las imágenes registradas, consiste en medir el incremento de energía en la banda de altas frecuencias espaciales generado por el desregistro en la imagen fundida. La aplicación de este criterio de evaluación permite asignar una medida de tipo geométrico a la calidad visual del registro de dos imágenes.
  - El segundo método de evaluación se basa en la determinación del campo vectorial de desplazamientos locales, y depende únicamente de la deformación de desregistro, por lo que puede utilizarse en aplicaciones de calibración.

En resumen, se ha resuelto el problema de registro de imágenes de bandas espectrales alejadas mediante el desarrollo de nuevos métodos de registro, cuya generalidad permite también su aplicación a tareas indirectamente relacionadas con el registro, como el reconocimiento de patrones, la creación de mosaicos, el seguimiento de blancos, o la integración de secuencias de imágenes. Asimismo, se han desarrollado métodos para la evaluación de los procedimientos empleados y se ha comprobado su funcionamiento mediante la aplicación de los métodos de registro desarrollados a la resolución de problemas reales de registro, en particular, registro de imágenes visibles, IR y SAR. La utilización de estos nuevos procedimientos de registro permite extender, en gran medida, el rango de aplicación de las técnicas de fusión de imágenes de bandas espectrales alejadas.





## Apéndice A: Modelo de transformación afín utilizado

Como modelo de transformación hemos elegido la descomposición:

$$T_m = \begin{pmatrix} \cos \alpha' & \sin \alpha' \\ -\sin \alpha' & \cos \alpha' \end{pmatrix} \cdot \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad (\text{A.1})$$

por la facilidad de interpretación de los parámetros elegidos y por la ortogonalidad que presentan entre sí. La principal diferencia con otras descripciones más tradicionales está casi únicamente en la elección del modelo para la transformación de sesgo, que aquí utiliza un único parámetro,  $k$ . En este modelo,  $k$  indica la desviación de la ortogonalidad, respecto del eje  $x$ , de las líneas paralelas al eje  $y$  en el espacio fuente.

A continuación mostraremos que este modelo es equivalente a la descomposición en valores singulares (SVD), que puede escribirse del modo siguiente:

$$T_{SVD} = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \cdot \begin{pmatrix} \cos \gamma & \sin \gamma \\ -\sin \gamma & \cos \gamma \end{pmatrix} \quad (\text{A.2})$$

Realizando productos y sustituyendo  $\alpha' \equiv \beta + \alpha$  en (A.1) se obtiene:

$$T_m = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \cdot \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \cdot \begin{pmatrix} a & bk \\ 0 & b \end{pmatrix} \quad (\text{A.3})$$

Comparando (A.2) y (A.3) puede verse que basta con demostrar la siguiente equivalencia:

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \cdot \begin{pmatrix} a & bk \\ 0 & b \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \cdot \begin{pmatrix} \cos \gamma & \sin \gamma \\ -\sin \gamma & \cos \gamma \end{pmatrix} \quad (\text{A.4})$$

Desarrollando e igualando elementos obtenemos el siguiente sistema de ecuaciones:

$$\begin{cases} a \cos \alpha = A \cos \gamma \\ -a \sin \alpha = -B \sin \gamma \\ bk \cos \alpha + b \sin \alpha = A \sin \gamma \\ -bk \sin \alpha + b \cos \alpha = B \cos \gamma \end{cases} \quad (\text{A.5})$$

De donde pueden deducirse las relaciones:

$$\operatorname{tg} \alpha = \frac{B}{A} \operatorname{tg} \gamma \quad (\text{A.6})$$

$$k = \frac{(A^2 - B^2) \operatorname{tg} \gamma}{AB(\operatorname{tg}^2 \gamma + 1)} \quad (\text{A.7})$$

En función de ellas pueden obtenerse expresiones para  $a$  y  $b$ :

$$b = \frac{A \operatorname{sen} \gamma}{k \cos \alpha + \operatorname{sen} \alpha} \quad (\text{A.8})$$

$$a = \frac{A}{B} (b - bk \operatorname{tg} \alpha) = b \left( \frac{A}{B} - k \operatorname{tg} \gamma \right) \quad (\text{A.9})$$

Para comprobar que estas relaciones son consistentes, deben analizarse todos los casos en los que pueden aparecer singularidades. En (A.6) y (A.7) no existe ningún problema, ya que ni  $A$  ni  $B$  pueden hacerse cero. En (A.7) tampoco hay problemas cuando la tangente de  $\gamma$  se hace infinita, en cuyo caso  $k$  tiende a cero.

En la expresión de  $b$  (A.8), sin embargo, hay dos casos que deben analizarse. El primero es en el que  $k$  y  $\operatorname{sen} \alpha$  valen los dos cero:

$$\lim_{\operatorname{sen} \alpha \rightarrow 0} \frac{A \operatorname{sen} \gamma}{\operatorname{sen} \alpha} = \lim_{\operatorname{sen} \alpha \rightarrow 0} \frac{A \operatorname{tg} \gamma \cos \gamma}{\operatorname{tg} \alpha \cos \alpha} = \lim_{\substack{\cos \alpha \rightarrow \pm 1 \\ \cos \gamma \rightarrow \pm 1}} \frac{A^2 \operatorname{tg} \gamma \cos \gamma}{B \operatorname{tg} \gamma \cos \alpha} = \pm \frac{A^2}{B} \quad (\text{A.10})$$

Y el segundo es en el que  $k \cos \alpha = -\operatorname{sen} \alpha$ . En este caso  $k = -\operatorname{tg} \alpha$ , que sustituido en (A.7) y teniendo en cuenta (A.6) permite llegar a la relación  $\operatorname{tg}^2 \alpha = -1$ , con lo que esta situación no puede darse.

Para (A.9) el único caso a tener en cuenta es en el que  $\operatorname{tg} \gamma$  se hace infinito. Sustituyendo la expresión de  $k$  (A.7) en el límite se obtiene:

$$\lim_{\operatorname{tg} \gamma \rightarrow \infty} k \operatorname{tg} \gamma = \lim_{\operatorname{tg} \gamma \rightarrow \infty} \frac{(B^2 - A^2) \operatorname{tg}^2 \gamma}{AB(\operatorname{tg}^2 \gamma - 1)} = \frac{(B^2 - A^2)}{AB} \quad (\text{A.11})$$

Con lo que queda demostrada la equivalencia del modelo elegido para las transformaciones lineales con la descomposición en valores singulares.

## Apéndice B: Reducción del ruido de barrido en imágenes IR

Las imágenes IR adquiridas por cámaras en las que la formación de la imagen se realiza mediante barridos de la escena sobre un detector suelen presentar saltos de intensidad en algunas líneas de la imagen, causados por una calibración imprecisa del detector. Ese incremento del nivel medio de gris se mantiene en algunas líneas y después desaparece, con lo que el efecto se manifiesta globalmente como bandas horizontales claras u oscuras, que dan lugar a segmentos inexistentes en la escena cuando se realiza la extracción de bordes. Para eliminar estas bandas es necesario corregir los incrementos en el nivel medio de gris de cada línea producidos por el ruido de barrido. Sin embargo, los contornos de los objetos presentes en la imagen también influyen sobre los incrementos de intensidad media de cada línea, y debe descontarse su valor para contabilizar únicamente la aportación debida al ruido de barrido.

La idea principal sobre la que se fundamenta el método desarrollado es que el nivel medio de gris en dos líneas consecutivas de la imagen, descontando los saltos de gradiente debidos a los contornos de los objetos de la escena, debe ser constante. En consecuencia, debe decidirse si las variaciones en el valor del gradiente son debidas a la presencia de objetos o al ruido de barrido, en cada píxel de una línea de la imagen. Para ello, es necesario clasificar los puntos de la imagen de grises como pertenecientes o no a los bordes de algún objeto. Aunque esto podría obligar a realizar una segmentación para la descomposición de la imagen en regiones uniformes, estudiando las imágenes en las que aparece este tipo de ruido hemos llegado a la conclusión de que la clasificación puede hacerse simplemente mediante la umbralización del gradiente.

Las medidas de la variación del nivel de gris en zonas uniformes de la imagen en las que existe un salto debido al ruido de barrido muestran que dicha variación es muy pequeña, del orden de 5 niveles de gris como máximo. De esta forma, todos aquellos puntos en los que la variación sea mayor que ese valor umbral, se considerarán como pertenecientes a bordes de objetos, y no se utilizarán para el cálculo del valor medio de la variación del nivel de gris.

Las medidas de la variación del nivel de gris no se realizan directamente, sino mediante la extracción del gradiente vertical de la imagen. Se ha utilizado el filtro de Sobel [9-13] para esta tarea, que consiste en convolucionar la imagen de entrada con la máscara siguiente:

-1	-2	-1
0	0	0
1	2	1

Como la máscara no está normalizada, el valor del gradiente asociado al paso de una región uniforme, de nivel de gris  $g_1$  a otra de nivel de gris  $g_2$  es el siguiente:

$$(1 + 2 + 1) \cdot g_2 - (1 + 2 + 1) \cdot g_1 = 4 \cdot (g_2 - g_1) \quad (\text{B.1})$$

El valor límite para el gradiente será entonces de  $4 \cdot 5 = 20$ , y todos los puntos de la imagen cuyo gradiente sea mayor no se tendrán en cuenta. La expresión para el gradiente medio de cada línea descontando aquellos puntos,  $\overline{\nabla I}(y_0)$ , será por tanto:

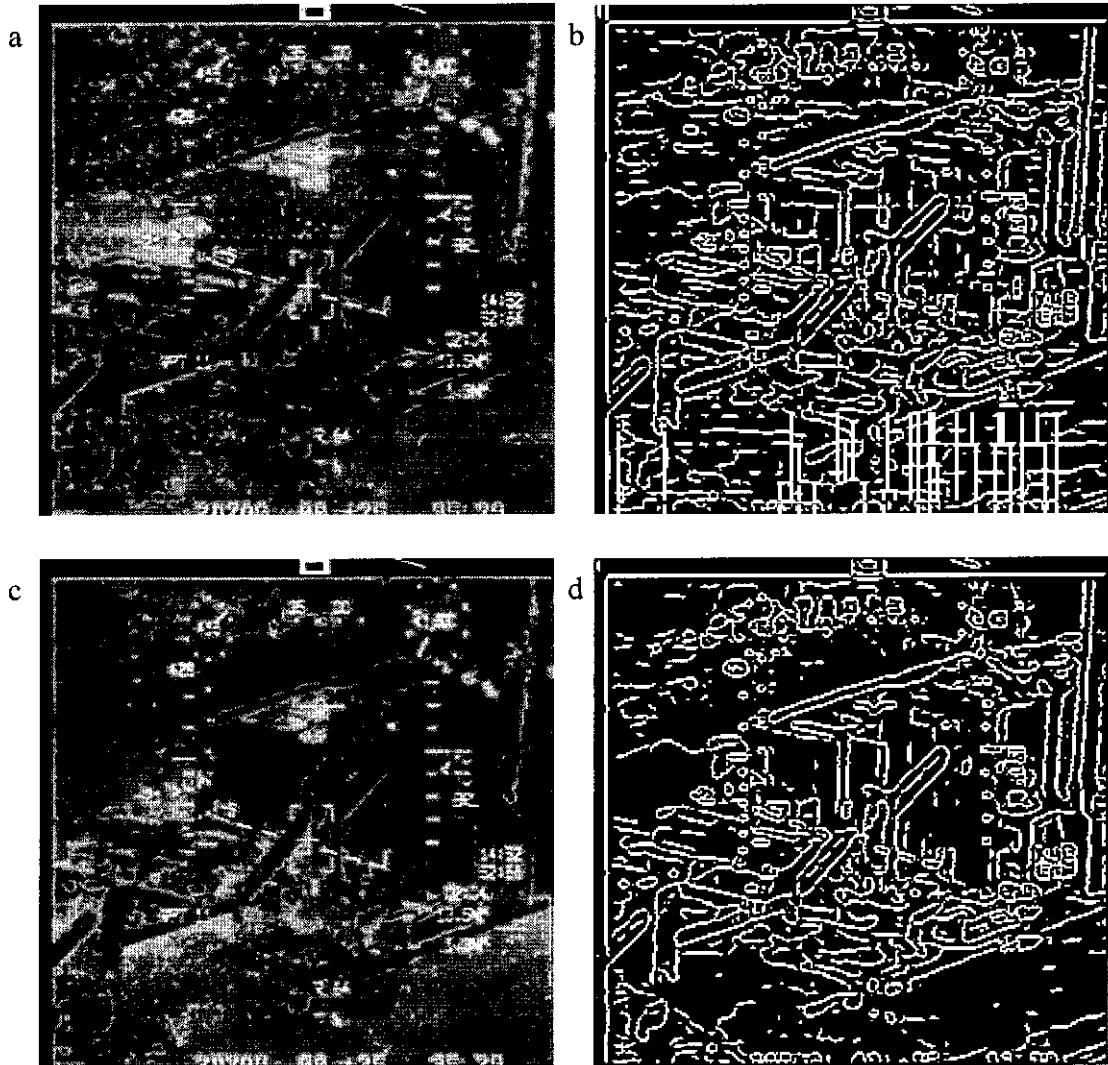
$$\overline{\nabla I}(y_0) = \frac{\sum_x \{ \nabla I(x, y_0) \mid \nabla I(x, y_0) \leq 20 \}}{\text{Card}\{ \nabla I(x, y_0) \leq 20 \}} \quad (\text{B.2})$$

donde  $I(x, y)$  es el nivel de gris del punto  $(x, y)$  de la imagen,  $\nabla I(x, y)$  indica el valor del gradiente en ese punto, y  $\text{Card}$  es el operador cardinal.

La corrección del ruido de barrido es, finalmente:

$$I(x_0, y_0) \leftarrow I(x_0, y_0) - \sum_{y=0}^{y_0} \overline{\nabla I}(y) \quad (\text{B.3})$$

En la figura B.1 se muestra un ejemplo de la utilización del filtro. La aplicación de un extractor de bordes (Canny,  $\sigma = 3.0$  píxeles) sobre una imagen IR con ruido de barrido (B.1a) genera múltiples bordes falsos (B.1b), que no pertenecen a la escena. La aplicación del método de filtrado desarrollado (B.1c), reduce en gran medida la cantidad de bordes espurios detectados (B.1d).



**Figura B.1.** Aplicación del filtro desarrollado. (a) Imagen IR con ruido de barrido y (b) bordes extraídos mediante Canny ( $\sigma = 3.0$  píxeles). (c) Resultado de la aplicación del filtro y (d) bordes resultantes.



## Apéndice C: Crecimiento Hexadecagonal de Regiones

### C.1. Introducción

Los modelos de crecimiento tienen muchas aplicaciones en el campo del tratamiento digital de imágenes, como el cálculo del área de influencia de determinados elementos presentes en las imágenes, extracción de esqueletos de figuras, generación de diagramas de Voronoi, etc. Sin embargo, la dificultad de desarrollar un algoritmo fácil de implementar y que consuma pocos recursos de proceso ha forzado la utilización de modelos de crecimiento bastante alejados del comportamiento isótropo.

Los métodos de crecimiento de formas binarias se han restringido tradicionalmente a la dilatación morfológica en regiones de  $3 \times 3$ , utilizando vecindades 4-conectadas u 8-conectadas [77]. Como las operaciones de dilatación y erosión suelen aplicarse sólo un reducido número de veces, la naturaleza anisótropa de los modelos de crecimiento en 4- y 8-vecinos no es crítica. Pero también es cierto que ese carácter anisótropo impide una utilización más extendida de este tipo de modelos.

Otros modelos deterministas para el crecimiento, como la difusión [78], crean patrones de crecimiento aproximadamente isótropos, pero son difíciles de implementar. Además, la discretización de las ecuaciones de difusión puede hacer que el algoritmo converja a estados estacionarios, a partir de los cuales no evolucione.

El modelo de distancia de Chamfer  $3/4$  [79] puede utilizarse también para la dilatación determinista de imágenes binarias, calculando el mapa de distancias para la imagen y realizando después una truncación del mismo hasta un cierto valor umbral. Esta técnica se fundamenta en el análisis local de regiones  $3 \times 3$  y crea patrones de crecimiento en forma de octógono irregular, que se aproximan mejor al crecimiento circular que los métodos de crecimiento basados en máscaras de 4 y 8 vecinos. Sin embargo, no es una técnica local, ya que toda la imagen debe ser procesada para simular el proceso de dilatación. La distancia de Chamfer  $5/7/11$  [79] obtiene mejores resultados, pero depende del análisis de regiones locales mayores ( $5 \times 5$ ). Otras técnicas globales [80] generan aproximaciones hexagonales y dodecagonales al crecimiento circular, pero sólo operan sobre rejillas hexagonales.



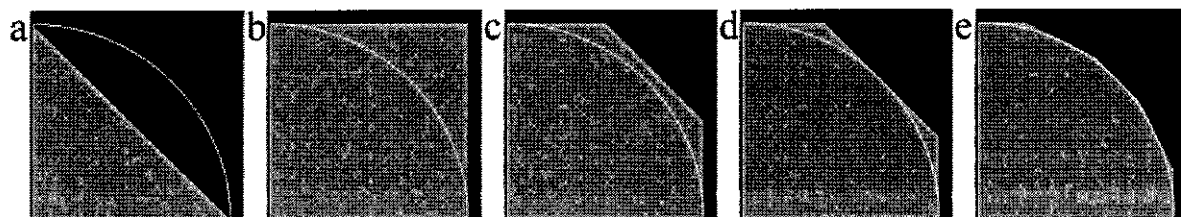
Los retardos temporales (*time-delays*) para el crecimiento de 8 vecinos [81] son modelos complejos que pueden generar formas circulares (entre otras), pero solamente después de un número determinado de iteraciones. Durante las etapas intermedias, sin embargo, la forma no es, en general, circular. Otros procedimientos generan distancias euclídeas sobre rejillas cuadradas [82], pero necesitan almacenar datos para los cálculos de distancia o para el mantenimiento de caminos de distancia mínima entre puntos.

En la referencia [83] se proponen diversos algoritmos para la propagación circular, y se presentan resultados empíricos de crecimiento hexadecagonal. Sin embargo, no se propone ningún modelo específico y, además, la técnica de redondeado de esquinas (*corner-rounding*) en la que se basan estos métodos de propagación evita que puedan ser aplicados al crecimiento de figuras.

Por todo ello se ha realizado un análisis detallado del crecimiento hexadecagonal, y se ha desarrollado un modelo sencillo para su aplicación. Sus resultados han sido, además, comparados con aquellos de los métodos de crecimiento 3x3 más habituales, así como con el modelo de crecimiento 5x5 más extendido, el Chamfer 5/7/11.

## C.2. Crecimiento Octogonal Regular

La aplicación iterada de los patrones de dilatación en 8-v (8 vecinos) o 4-v (4 vecinos) a un píxel aislado, genera bien un cuadrado, en el caso del crecimiento en 8-v, o un rombo, en el caso del crecimiento en 4-v (véase la figura C.1).



**Figura C.1.** Aplicación de los diferentes modelos de crecimiento a un píxel aislado situado en la esquina inferior izquierda de las imágenes, y el arco circular ideal para una circunferencia de radio 255 píxeles. (a) Crecimiento en 4-v, (b) 8-v, (c) crecimiento octogonal (iteración de 4-v, 8-v), (d) crecimiento octogonal regular y (e) crecimiento hexadecagonal propuesto. Sólo se muestran los primeros cuadrantes.

En efecto, iniciando el proceso de crecimiento desde un píxel aislado situado en el origen de coordenadas,  $O = (0, 0)$ , y después de  $n$  iteraciones del modelo de crecimiento en 4-v, el conjunto de píxeles contenidos en la figura de dilatación es el siguiente:

$$C_4(n) = \{(x, y) \in Z^2 : |x| + |y| \leq n\} \quad (C.1)$$

donde  $x$  e  $y$  son enteros y  $Z$  es el conjunto de los números enteros. La figura resultante es un rombo, de vértices  $(n, 0)$ ,  $(0, n)$ ,  $(-n, 0)$  y  $(0, -n)$ .

En el caso del crecimiento en 8-v, el conjunto es:

$$C_8(n) = \{(x, y) \in Z^2 : |x| \leq n, |y| \leq n\} \quad (C.2)$$

que se corresponde con un cuadrado de vértices  $(n, n)$ ,  $(-n, n)$ ,  $(-n, -n)$  y  $(n, -n)$ .

Una mejor aproximación al modelo isótropo, que crece como un octógono irregular (figura 1.c) puede obtenerse fácilmente mediante la combinación de los crecimientos en 4-v y 8-v, aplicando alternativamente uno u otro [77]. Y, puesto que el área de  $C_4(1)$  es una mejor aproximación al área del círculo discreto de radio 1 que el área de  $C_8(1)$ , el crecimiento en 4-v debe aplicarse primero [84]. Después de  $n$  iteraciones resulta el siguiente conjunto:

$$C_{4,8}(n) = \left\{ (x, y) \in Z^2 : |x| \leq n, |y| \leq n, |x| + |y| \leq \left\lfloor \frac{3}{2}n \right\rfloor \right\} \quad (C.3)$$

donde  $\lfloor x \rfloor$  significa el número entero más cercano y no mayor que  $x$ .

Las dimensiones del octógono resultante, no obstante, dependen sólo del número de veces que se aplica cada uno de los modelos, pero no de su orden de aplicación, como se detalla en [83]. Entonces, si los crecimientos en 4-v y 8-v se aplican  $n_4$  y  $n_8$  veces, respectivamente, el conjunto de puntos que pertenecen a la figura dilatada es:

$$C'_{4,8}(n_4, n_8) = \{(x, y) \in Z^2 : |x| \leq n, |y| \leq n, |x| + |y| \leq n_4 + 2n_8\} \quad (C.4)$$

donde  $n = n_4 + n_8$  es el número total de iteraciones. Los vértices de este octógono son  $(n, \pm n_8)$ ,  $(n_8, \pm n)$ ,  $(-n_8, \pm n)$  y  $(-n, \pm n_8)$ . Para  $n_4 = 0$ , el octógono degenera en un cuadrado, y para  $n_8 = 0$  en un rombo, en concordancia con los modelos de 8 y 4 vecinos.

Esta aproximación octogonal puede mejorarse haciendo que todos los lados del octógono sean tangentes a la circunferencia, lo cual generaría un octógono regular. Para conseguir esto, debe llegarse a un equilibrio entre los crecimientos en 4 y 8 vecinos.

En el caso de un octógono regular circunscrito a un círculo de radio  $n$ , la posición del vértice superior situado más a la derecha es  $(n, n \cdot (\sqrt{2} - 1))$ . Para el octógono  $C'_{4,8}$  ese vértice se encuentra en  $(n, n_8)$ , por lo que, para ser regular, debe cumplir  $n_8 = n \cdot (\sqrt{2} - 1)$  ó  $n_8 / n = \sqrt{2} - 1$ .

Por lo tanto, para decidir si debe aplicarse un crecimiento en 4-v o en 8-v, la fracción  $n_8/n$  debe evaluarse en cada iteración del algoritmo. Si su valor es mayor que  $\sqrt{2} - 1$ , se aplica crecimiento en 4-v; en caso contrario se utiliza 8-v. Y, como se ha explicado anteriormente, debe utilizarse 4-v para la primera iteración.

Para un pequeño número de iteraciones, el valor de  $n_8 / n$  puede aproximarse por:

$$\frac{n_8}{n} = \sqrt{2} - 1 \approx \frac{1}{2} - \frac{1}{12} \quad (C.5)$$

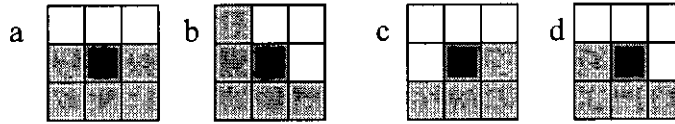
Entonces, aplicando crecimiento en 8-v para aquellas iteraciones que sean múltiplos de 2, pero no de 12, se obtiene un octógono regular. Esta regla de control es válida para círculos de hasta 410 píxeles de radio. Pueden obtenerse mejores aproximaciones añadiendo términos a la serie de aproximación a la fracción (el siguiente termino es  $-1/410$ ).

Los resultados de este modelo de crecimiento pueden verse en la figura C.1d.

### C.3. Modelo para el crecimiento hexadecagonal

Los vértices del octógono regular son los puntos que más se desvían de la circunferencia ideal. Sería muy útil, por tanto, controlar de alguna forma la velocidad de crecimiento en las direcciones de esos vértices. En ese sentido, es interesante el hecho de que, a partir de la tercera iteración del algoritmo de crecimiento octogonal regular, cada punto de los bordes del octógono puede clasificarse por su vecindad 3x3 como uno de los dos tipos siguientes:

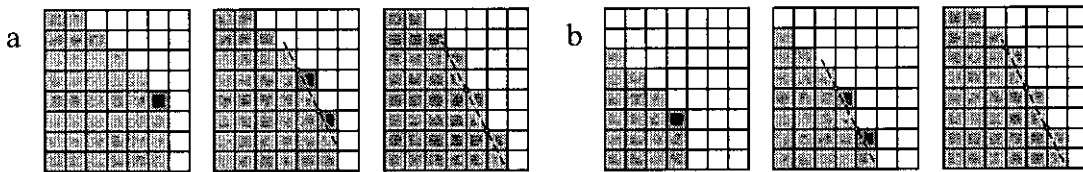
- *píxeles lado*: cinco de sus vecinos pertenecen al octógono. Pueden ser identificados por las máscaras de las figuras C.2a y C.2b, y sus cuatro posibles rotaciones de 90 grados.
- *píxeles vértice*: sólo cuatro de sus vecinos pertenecen al octógono. Pueden identificarse mediante las máscaras de las figuras C.2c y C.2d, y sus rotaciones.



**Figura C.2.** Tipos de vecindades para los píxeles del borde del octógono. (a), (b) Máscaras para los píxeles lado. (c), (d) Máscaras para los píxeles vértice.

Los píxeles vértice son, por tanto, fáciles de identificar, bien comprobando que sólo tienen 4 vecinos que pertenecen al octógono o mediante las máscaras 3x3 adecuadas. También pueden utilizarse técnicas de análisis de contornos más complejas, como las descritas en [85].

Una forma de controlar la velocidad de crecimiento de los píxeles vértice consiste en borrarlos en algunas iteraciones. A esta técnica, descrita en [83], se la denomina redondeado de esquinas (*corner-rounding*). Como se muestra en la figura C.3a, la aplicación iterada de la eliminación de vértices equivale a la truncación del vértice del octógono mediante un nuevo lado de pendiente  $-2$  (ó  $2$ ,  $-1/2$ ,  $1/2$ , según el vértice).



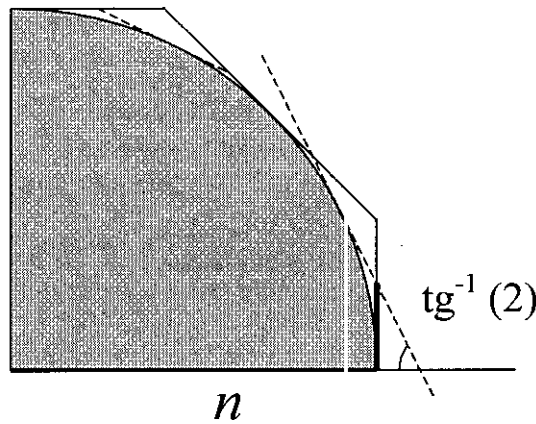
**Figura C.3.** (a) Resultado de la eliminación de vértices (2 iteraciones). Los píxeles eliminados se muestran en color gris oscuro. (b) Resultado de la inhibición del crecimiento de los píxeles vértice (2 iteraciones). Los píxeles que no crecieron en la iteración correspondiente se muestran en gris oscuro. El resultado final es equivalente al de la figura C.3a.

Uno de los inconvenientes de la supresión de vértices es que el algoritmo necesita realizar dos pasadas en cada iteración: una para el crecimiento y otra para la supresión. Además, cuando el algoritmo de dilatación se aplica a una forma binaria genérica, aparecen problemas en las zonas donde interfieren dos formas crecientes [83], puesto que los puntos vértice no pueden ser identificados allí mediante máscaras.

Este problema puede evitarse si en lugar de la eliminación de vértices se inhibe su crecimiento en algunas iteraciones [81]. Como se muestra en la figura 3, la inhibición del crecimiento de los píxeles vértice es equivalente a su eliminación, lo que puede comprobarse mediante la aplicación iterada de los crecimientos en 4 y 8 vecinos a las configuraciones de lado y vértice de la figura 2.

Para regular la inhibición del crecimiento de los vértices se ha desarrollado una regla de control similar a la utilizada para el crecimiento octogonal regular. Al igual que entonces, es necesario determinar las coordenadas del vértice superior situado más a la derecha.

Una etapa de supresión (o de inhibición) mueve un píxel hacia abajo el vértice superior situado más a la derecha. Por tanto, para una cierta iteración,  $n$ , la técnica de supresión o inhibición deberá ser aplicada tantas veces como la distancia vertical existente entre el vértice superior situado más a la derecha del octógono y el vértice correspondiente del hexadecágono resultante de la truncación del octógono regular con una recta, tangente a la circunferencia de radio  $n$  y de pendiente  $-2$  (figura C.4).



**Figura C.4.** Truncación de un octógono regular circunscrito a una circunferencia de radio  $n$  mediante una recta de pendiente  $-2$ .

Como el vértice superior situado más a la derecha en el octógono truncado tiene por coordenadas  $(n, n \cdot (\sqrt{5} - 2))$ , el número de supresiones o inhibiciones necesarias,  $n_s$ , es  $n_s = n \cdot (\sqrt{2} - 1) - n \cdot (\sqrt{5} - 2) = n \cdot (1 + \sqrt{2} - \sqrt{5})$ , ó  $n_s / n = 1 + \sqrt{2} - \sqrt{5}$ . Este valor puede aproximarse por:

$$\frac{n_s}{n} = 1 + \sqrt{2} - \sqrt{5} \approx \frac{1}{5} - \frac{1}{45} \quad (\text{C.6})$$

que es válido para circunferencias de radio menor que 2718 píxeles. Por lo tanto, la inhibición en el crecimiento de los vértices debe aplicarse en aquellas iteraciones que sean múltiplo de 5 pero no de 45. La aplicación de esta corrección genera un crecimiento en forma de polígono de 16 lados, o hexadecágono (figura C.1e). Esta regla de control también es válida para la técnica del redondeo de esquinas, y puede verse como un fundamento teórico de los resultados empíricos presentados en la referencia [83].

La principal ventaja de la técnica desarrollada es que la verificación de la condición de vértice puede hacerse simplemente contando el número de vecinos del píxel, en lugar de compararlo con las 8 máscaras posibles. La razón es que las configuraciones en las que el crecimiento debería ser inhibido, pero no cumplen la condición de tener sólo 4 vecinos, evolucionan exactamente hacia la misma configuración independientemente de si se realiza o no la inhibición del crecimiento (tanto para etapas de crecimiento en 8-v como en 4-v). Y puesto que no se necesita un proceso de comparación con máscaras, se evitan los problemas de filtrado relacionados con la técnica del redondeo de esquinas (como en la referencia [83]).

#### C.4. Resultados y aplicaciones

La calidad de cada una de las aproximaciones al crecimiento circular puede evaluarse mediante la comparación de un círculo con las figuras resultantes de la aplicación iterada de los diferentes modelos de crecimiento a un píxel aislado. La tabla C.1 muestra los errores de área, distancia y los cocientes entre el área y el cuadrado del radio para cada uno de los modelos. El error en el área se calcula de la siguiente manera:

$$Error\_area = 100 \cdot \left| \frac{A_m - A_c}{A_c} \right| \quad (C.7)$$

donde  $A_m$  es el área del polígono asociado al modelo de crecimiento, y  $A_c$  es el área del círculo inscrito en el polígono. El error en la distancia se calcula de forma similar:

$$Error\_distancia = 100 \cdot \left| \frac{r_c' - r_c}{r_c} \right| \quad (C.8)$$

donde  $r_c$  y  $r_c'$  son los radios de las circunferencias inscrita y circunscrita. Cada modelo de crecimiento asigna una misma distancia (igual al radio de la circunferencia inscrita) a todos

los puntos del contorno del polígono. Por lo tanto, el mayor error en la medida de la distancia se corresponderá con el del vértice del polígono que más se aleja de la circunferencia inscrita. Y, por definición, la distancia euclídea desde el centro de crecimiento a ese vértice es el radio de la circunferencia circunscrita al polígono.

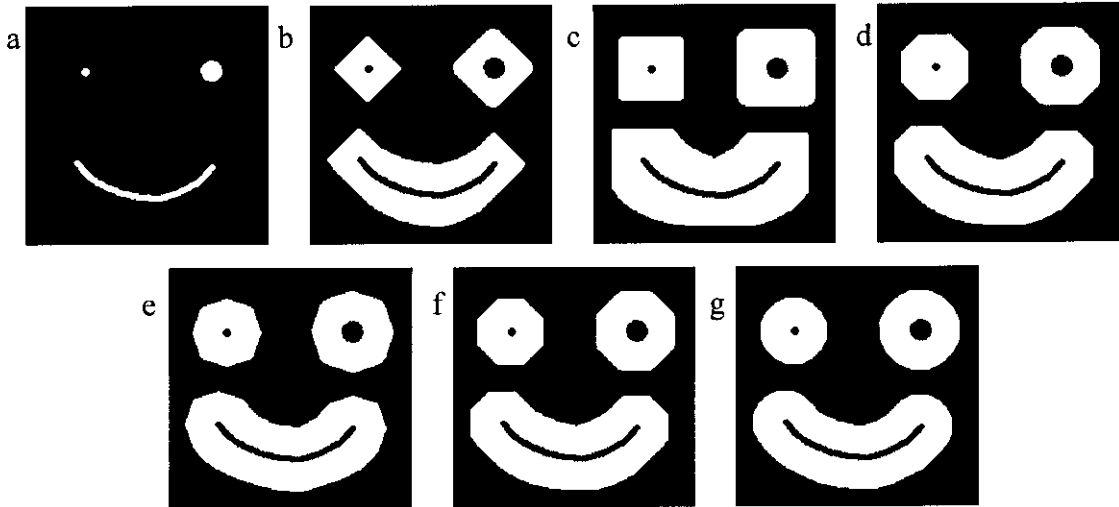
Para la realización de las comparaciones debe hacerse notar que el modelo de 4-v crece como un polígono inscrito al crecimiento circular ideal, al contrario que el resto de los modelos de crecimiento. Esta distinción se ha hecho en concordancia con el modo usual de aplicación del algoritmo de crecimiento en 4 vecinos. También hay que mencionar que los modelos de crecimiento basados en las distancias de Chamfer 3/4 y 5/7/11 han sido normalizados mediante  $\sqrt{10}$  y  $\sqrt{26}$ , respectivamente, en lugar de los coeficientes de normalización usuales (3 y 5), para que generen crecimientos poligonales circunscritos a la circunferencia de radio correspondiente.

Los resultados de la tabla C.1 muestran una clara superioridad del método desarrollado sobre los demás modelos de crecimiento basados en vecindades 3x3, e incluso que la simulación de dilatación basada en la distancia de Chamfer 5/7/11, que trabaja sobre vecindades 5x5.

En la figura C.5 se muestra una comparación de los resultados de la aplicación de cada uno de los modelos de crecimiento 3x3 analizados para la dilatación de una imagen binaria. Como puede verse, el modelo hexadecagonal propuesto tiene prestaciones superiores a las de los demás.

Modelo de crecimiento	Cociente entre el área y $n^2$	% error de área	% error de distancia
4 vecinos (inscrito)	2	36.34%	70.71%
8 vecinos	4	27.32%	29.29%
Octogonal	3.5	11.41%	10.56%
Distancia de Chamfer 3/4	3.33333...	6.10%	10.56%
Octogonal regular	3.31371...	5.48%	7.61%
Distancia de Chamfer 5/7/11	3.24156...	3.18%	3.52%
Hexadecagonal (propuesto)	3.18677...	1.44%	2.67%
Crecimiento circular ideal	$\pi = 3.14159...$	0%	0%

**Tabla C.1.** Cocientes entre el área y el cuadrado del radio, errores de área y errores de distancia para cada uno de los modelos de crecimiento analizados.



**Figura C.5.** Comparación de los diferentes modelos de dilatación 3x3 analizados, después de 24 iteraciones. (a) Imagen binaria con algunas formas aisladas. (b) Dilatación mediante 4-v. (c) Dilatación por 8-v. (d) Dilatación por crecimiento octogonal (4,8). (e) Simulación de dilatación mediante distancia Chamfer 3/4. (f) Dilatación por crecimiento octogonal regular, y (g) modelo de crecimiento hexadecagonal propuesto.

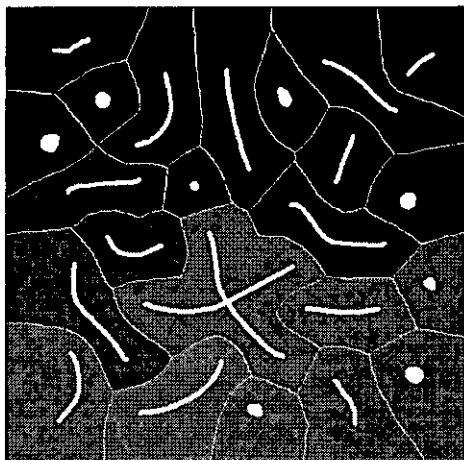
Además de las aplicaciones más evidentes, como la dilatación y la erosión, o la extracción de esqueletos, existen otras áreas de aplicación donde los algoritmos de crecimiento tradicionales no podían utilizarse, debido a sus limitaciones. Es el caso de la generación de diagramas de Voronoi a partir de un conjunto de píxeles aislados [64, 66, 67], o a partir de formas arbitrarias [65, 68]. El diagrama de Voronoi de un conjunto de elementos divide el espacio en un conjunto de regiones, cada una de las cuales contiene los puntos del espacio más cercanos a uno de los elementos que a los demás. La generación de los diagramas de Voronoi es un problema analítico complejo, pero que puede resolverse fácilmente mediante el etiquetado de los elementos y su posterior expansión mediante un algoritmo de crecimiento, como el modelo hexadecagonal propuesto. Más aún, el grafo de conectividad de los elementos (la triangulación de Delaunay asociada) se recupera simplemente por la comparación de las etiquetas de regiones adyacentes, creando una nueva arista del grafo por cada frontera del diagrama.

La figura C.6 muestra el resultado de la aplicación del algoritmo para la extracción del diagrama de Voronoi de un conjunto de figuras binarias arbitrarias.

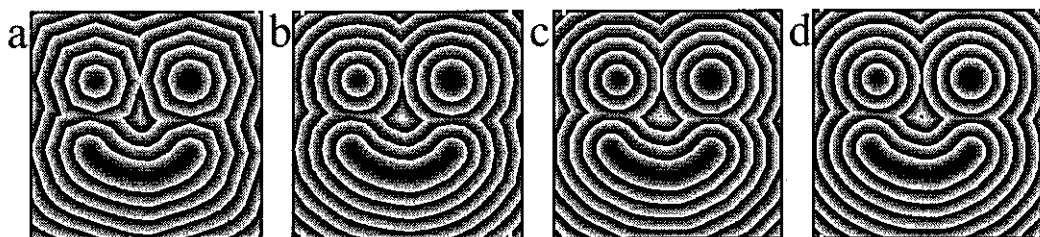
Algunas operaciones, como la extracción de esqueletos [86] o la determinación de agregaciones de píxeles [10], se basan en el cálculo de mapas de distancias, que pueden obtenerse a partir de un método de crecimiento mediante el etiquetado de los píxeles con el número de la iteración en la que son rellenados. El modelo más utilizado para la obtención



de estos mapas suele ser el Chamfer 3/4, o el Chamfer 5/7/11 si se necesita más precisión. Sin embargo, como demuestra la tabla C.1, la aproximación del método hexadecagonal al crecimiento circular es mejor que la de aquellos. En la figura C.7 se comparan los resultados de la generación del mapa de distancias para los segmentos de la imagen de la figura C.5a, producidos por estos tres modelos.



**Figura C.6.** Diagrama de Voronoi de un conjunto de figuras binarias arbitrarias, obtenido mediante el modelo de crecimiento hexadecagonal propuesto.



**Figura C.7.** Mapas de distancia para los segmentos de la figura C.5a. (a) Distancia de Chamfer 3/4, (b) distancia de Chamfer 5/7/11, (c) distancia hexadecagonal utilizando el modelo propuesto y (d) discretización de la distancia euclídea. Para apreciar mejor el patrón de crecimiento se muestran las distancias módulo 32. Las imágenes son de 256x256 píxeles y los errores máximos para cada uno de los modelos son aquí de 12.39 píxeles para el Chamfer 3/4, 3.44 para el Chamfer 5/7/11 y 2.16 para el crecimiento hexadecagonal propuesto.

Si los modelos de crecimiento van a utilizarse como medida de distancia, es posible refinarlos mediante el ajuste de un coeficiente de normalización, que minimice la distancia del polígono de crecimiento a la circunferencia ideal, en lugar de quedar inscrito o circunscrito a la misma. Esta técnica, descrita en [87] y [88] para su aplicación al modelo de Chamfer 5/7/11, puede utilizarse también para refinar el método de crecimiento hexadecagonal.

## Bibliografía

1. J. M. Lloyd, *Thermal Imaging Systems*, Plenum Press, New York (1975).
2. F. T. Ulaby, *Microwave Remote Sensing*, R. K. Moore, A. K. Fung, Artech House (1982).
3. J. Santamaría, M. T. Gomez, "Visible-IR Image Fusión Based on Gabor Wavelets Decomposition", *EOS Annual Meetings Digest* **3**, 97-98 (1993).
4. S. R. F. Simps, M. A. Phillips, "Target signature consistency of image data fusión alternatives", *Optical Engineering* **36**(3), 743-754 (1997).
5. H. Li, B. S. Manjunath, S. K. Mitra, "Multisensor Image Fusión Using the Wavelet Transform", *Graph. Mod. Img. Proc.* **57**(3), 235-245 (1995).
6. D. A. Yocky, "Artifacts in wavelet image merging", *Optical Engineering* **35**(7), 2094-2101 (1996).
7. O. Nestares, R. Navarro, J. Portilla, A. Tabernero, "Efficient spatial-domain implementation of a multiscale image representation based on Gabor functions", *Journal of Electronic Imaging* **7**(1), 166-173 (1998).
8. L. G. Brown, "A Survey of Image Registration Techniques", *ACM Computing Surveys* **24**(4), 325-376 (1992).
9. S.-T. Bow, *Pattern Recognition and Image Preprocessing*, Marcel Dekker, New York (1992).
10. R. M. Haralick and L. Shapiro, *Computer and Robot Vision*, Addison-Wesley, Reading, Massachusetts (1992).
11. W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, (1991).
12. R. C. Gonzalez, P. Wintz, *Digital Image Processing. Second Edition*, Addison Wesley (1987).
13. B. Jähne, *Digital Image Processing, Concepts, Algorithms and Scientific Applications*. Springer-Verlag, Heidelberg (1991).
14. P. Aschwandén, W. Guggenbuehl, "Experimental Results from a Comparative Study on Correlation-type Registration Algorithms", en *Robust Computer Vision*, Foerstner and Ruwiedel, Wichmann, 268-289 (1993).
15. G. D. Hager, P. N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination", *IEEE PAMI* **20**(10), 1025-1039 (1998).
16. A. Domingo, J. Santamaría, M. T. Gómez, "Algoritmo de Seguimiento Automático de Objetos mediante Correlación de Imágenes", *Optica Pura y Aplicada* **26**, 613-624 (1993).

17. R. K. Wang, C. R. Chatwin, M. Y. Huang, "Correlation Filter with a Noise Adaptive Discriminant Capability", *Optik* **99**(1), 18-24 (1995).
18. M. S. Alam, J. G. Bogner, S. Cain, B. J. Yasuda, "Fast Registration and Reconstruction of Aliased Low-Resolution Frames by Use of a Modified Maximum-Likelihood Approach", *Applied Optics* **37**(8), 1319-1328 (1998).
19. J. D. Dunlop, G. H. Holder, E. F. LeDrew, "Image Matching using Spatial Frequency Signatures", *Proceedings of the IGARSS'89 and Canadian Symposium on Remote Sensing*, Vol. **3**, 1273-1276 (1989).
20. R. Venkateswarlu, B. N. Chatterjee, "Analysis of Image Registration Algorithms for Infrared Images", *SPIE* Vol. **1699**, 442-451 (1992).
21. P. Nesi, A. Del Bimbo, D. Ben-Tzvi, "A Robust Algorithm for Optical Flow Estimation", *Computer Vision and Image Understanding* **62**(1), 59-68 (1995).
22. C.-S. Fuh, P. Maragos, "Motion Displacement Estimation using an Affine Model for Image Matching", *Optical Engineering* **30**(7), 881-887 (1991).
23. T. Y. Tian, M. Shah, "Recovering 3D Motion of Multiple Objects using Adaptive Hough Transform", *IEEE PAMI* **19**(10), 1178-1183 (1997).
24. M. E. Spetsakis, "Optical Flow Estimation using Discontinuity Conforming Filters", *Computer Vision and Image Understanding* **68**(3), 276-289 (1997).
25. B. S. Reddy, B. N. Chatterji, "An FFT-Based Technique for Translation, Rotation and Scale-Invariant Image Registration", *IEEE Transactions on Image Processing* **5**(8), 1266-1271 (1996).
26. J. Sato, R. Cipolla, "Image Registration using Multi-Scale Texture Moments", *Image and Vision Computing* **13**(5), 341-353 (1995).
27. J. Flusser, T. Suk, "Degraded Image Analysis: An Invariant Approach", *IEEE PAMI* **20**(6), 590-603 (1998).
28. Q. Zheng, R. Chellapa, "A Computational Vision Approach to Image Registration", *IEEE Transactions on Image Processing* **2**(3), 311-326 (1993).
29. Y. Zhang, J. J. Gerbrands, "Method for Matching General Stereo Planar Curves", *Image and Vision Computing* **13**(8), 645-655 (1995).
30. A. K. Jain, Y. Zhong, S. Lakshmanan, "Object Matching Using Deformable Templates", *IEEE PAMI* **18**(3), 267-278 (1996).
31. D. Allison, M. J. A. Zemerly, J.-P. Muller, "Automatic Seed Point Generation for Stereo Matching and Multi-Image Registration", *IGARSS'91 Proceedings of the 11<sup>th</sup> Annual International Geoscience and Remote Sensing Symposium*, Vol. **4**, 2417-2421 (1991).
32. H. H. Li, Y.-T. Zhou, "Automatic visual/IR image registration", *Optical Engineering* **35**(2), 391-400 (1996).

33. G. Medioni, R. Nevatia, "Segment-based Stereo Matching", *Computer Vision Graphics and Image Processing* **31**, 2-18 (1985).
34. Z. Zhang, "Estimating Motion and Structure from Correspondences of Line Segments between Two Perspective Images", *IEEE PAMI* **17**(12), 1129-1139 (1995).
35. P. Gros, O. Bournez, E. Boyer, "Using Local Planar Geometric Invariants to Match and Model Images of Line Segments", *Computer Vision and Image Understanding* **69**(2), 135-155 (1998).
36. S. Sull, N. Ahuja, "Integrated Matching and Segmentation of Multiple Features in Two Views", *Computer Vision and Image Understanding* **62**(3), 279-297 (1995).
37. C. J. Taylor, D. J. Kriegman, "Structure and Motion from Line Segments in Multiple Images", *IEEE PAMI* **17**(11), 1021-1032 (1995).
38. R. Horaud, T. Skordas, "Stereo Correspondence Through Feature Grouping and Maximal Cliques", *IEEE PAMI* **11**(11), 1168-1180 (1989).
39. D. P. Huttenlocher, S. Ullman, "Recognizing Solid Objects by Alignment with an Image", *International Journal of Computer Vision* **5**(2), 195-212 (1990).
40. Behzad Kamgar-Parsi, Behrooz Kamgar-Parsi, "Matching Sets of 3D Line Segments with Application to Polygonal Arc Matching", *IEEE PAMI* **19**(10), 1090-1099 (1997).
41. S. O. Mason, K. W. Wong, "Image Alignment by Line Triples", *Photogrammetric Engineering and Remote Sensing* **58**(9), 1329-1334 (1992).
42. J.-W. Hsieh, H.-Y. M. Liao, K.-C. Fan, M.-T. Ko, Y.-P. Hung, "Image Registration Using a New Edge-Based Approach", *Computer Vision and Image Understanding* **67**(2), 112-130 (1997).
43. N. Sang, T. Zhang, "Rotation and Scale Change Invariant Point Pattern Relaxation Matching by the Hopfield Neural Network", *Optical Engineering* **36**(12), 3378-3385 (1997).
44. W.-H. Wang, Y.-C. Chen, "Image Registration by Control Points Pairing Using the Invariant Properties of Line Segments", *Pattern Recognition Letters* **18**, 269-281 (1997).
45. G. Z. Yang, P. Burger, D. N. Firmin, S. R. Underwood, "Structure adaptative anisotropic image filtering", *Image and Vision Computing* **14**, 135-145 (1996).
46. C. H. Li, C. K. Lee, "Image Smoothing Using Parametric Relaxation", *Graphical Models and Image Processing* **57**(2), 161-174 (1995).
47. J. F. Canny, "A Computational Approach to Edge Detection", *IEEE PAMI* **8**(6), 679-698 (1986).

48. T.-H. Yu, S. K. Mitra, "Efficient Approach for the Detection of Diffuse Edges", *Optical Engineering* **35**(12), 3522-3530 (1996).
49. E. Coiras, J. Santamaría, C. Miravet, "Hexadecagonal region growing", *Pattern Recognition Letters* **19**, 1111-1117 (1998).
50. J. D. Foley, A. van Dam, S. K. Feiner, J. F. Hughes, R. L. Phillips, *Computer Graphics: Principles and Practice*, Addison-Wesley (1982).
51. L. N. Kanal, B. A. Lambird, D. Lavine, G. C. Stockman, "Digital Registration of Images from Similar and Dissimilar Sensors", *Proceedings of the International Conference on Cybernetics and Society*, 347-351 (1981).
52. I. Fermin, A. Imiya, "Planar Motion Detection by Randomized Triangle Matching", *Pattern Recognition Letters* **18**, 741-749 (1997).
53. J. Sklansky, "On the Hough Technique for Curve Detection", *IEEE Transactions on Computers* **27**(10), 923-926 (1978).
54. H. Kälviäinen, P. Hirvonen, L. Xu, E. Oja, "Probabilistic and non-probabilistic Hough transforms: overview and comparisons", *Image and Vision Computing* **13**(4), 239-252 (1995).
55. P. S. Wu, M. Li, "Pyramid adaptative dynamic Hough transform to detect edges with arbitrary shapes", *Optical Engineering* **36**(5), 1425-1430 (1997).
56. D. Shaked, O. Yaron, N. Kiryati, "Deriving Stopping Rules for the Probabilistic Hough Transform by Sequential Analysis", *Computer Vision and Image Understanding* **63**(3), 512-526 (1996).
57. C. F. Olson, "Probabilistic Indexing for Object Recognition", *IEEE PAMI* **17**(5), 518-522 (1995).
58. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C. Second Edition*, Cambridge University Press (1994).
59. R. Navarro, A. Taberner, "Gaussian Wavelet Transform: Two Alternative Fast Implementations for Images", *Multidimensional Systems and Signal Processing* **2**, 421-436 (1991).
60. N. B. Nill, "Scene power spectra: the moment as an image quality merit factor", *Applied Optics* **15**(11), 2846-2854 (1976).
61. J. A. Saghri, P. S. Cheatham, "Image quality measure based on a human visual system model", *Optical Engineering* **28**(7), 813-818 (1989).
62. N. B. Nill, B. H. Bouzas, "Objective image quality measure derived from digital image power spectra", *Optical Engineering* **31**(4), 813-825 (1992).
63. H. H. Barrett, C. K. Abbey, E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions", *J. Opt. Soc. Am.* **15**(6), 1520-1535 (1998).

64. V.J.D. Tsai, "Fast Topological Construction of Delaunay Triangulations and Voronoi Diagrams", *Computer and Geosciences*, **19**(10), 1463-1474 (1993).
65. J.J. Chou, "Voronoi Diagrams for Planar Shapes", *IEEE Computer Graphics and Applications*, 52-59 (March, 1995).
66. Y. Oishi, K. Sugihara, "Topology-Oriented Divide-and-Conquer Algorithm for Voronoi Diagrams", *Graph. Mod. Img. Proc.* **57**(4), 303-314 (1995).
67. R. E. Sequeira, F. J. Prêteux, "Discrete Voronoi Diagrams and the SKIZ Operator: A Dynamic Algorithm", *IEEE PAMI* **19**(10), 1165-1170 (1997).
68. W. Guan, S. Ma, "A List-Processing Approach to Compute Voronoi Diagrams and the Euclidean Distance Transform", *IEEE PAMI* **20**(7), 757-761 (1998).
69. C. J. Oliver, "Synthetic-aperture radar imaging", *J. Phys. D: Appl. Phys.* **22**, 871-890 (1989).
70. C. J. Oliver, "Information from SAR images", *J. Phys. D: Appl. Phys.* **24**, 1493-1514 (1991).
71. C. J. Oliver, D. Blacknell, R. G. White, "Optimum edge detection in SAR", *IEEE Proc.-Radar. Sonar Navig.* **143**(1), 31-40 (1996).
72. K. M. Nam, D. J. Park, R.-H. Park, "Multiresolution edge detection in speckle imagery", *Optical Engineering* **34**(3), 922-932 (1995).
73. H. Ossman, S. D. Blostein, "SAR imagery segmentation using probabilistic winner-take-all clustering", *SPIE* **2757**, 217-226 (1996).
74. S. M. Ali, R. E. Burge, "New Automatic Techniques for Smoothing and Segmenting SAR Images", *Signal Processing* **14**(4), 335-346 (1988).
75. C. Miravet, J. Santamaría, E. Coiras, J. Ureña, "Generación Automática de Mosaicos. Aplicación de Técnicas de Fusión de Imágenes", *Revista de Teledetección* **10**, 31-38 (1998).
76. C. Miravet, E. Coiras, J. Santamaría, "Vectorización semi-automática de redes lineales en mapas cartográficos en color", *Revista de Teledetección* **10**, 49-62 (1998).
77. A. Rosenfeld and J.L.Pfaltz, "Distance Functions on Digital Pictures", *Pattern Recognition* **1**, 33-61 (1968).
78. S. Thompson and A. Rosenfeld, "Isotropic Growth on a Grid", *Pattern Recognition* **28**(2), 241-253 (1995).
79. G. Borgefors, "Distance Transformation in Digital Images", *Computer Vision Graphics and Image Processing* **34**, 344-371 (1986).
80. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London (1982).

81. S. Thompson and A. Rosenfeld, "Growth Processes Based on 8-Neighbor Time Delays", *Pattern Recognition* **30**(2), 321-337 (1997).
82. P.E. Danielsson, "Euclidean Distance Mapping", *Comp. Graph. Img. Proc.* **14**, 227-248 (1980).
83. Z. Kulpa, B. Kruse, "Algorithms for Circular Propagation in Discrete Images", *Computer Vision Graphics and Image Processing* **24**, 305-328 (1983).
84. P. P. Das, "Best Simple Octagonal Distances in Digital Geometry", *Journal of Approximation Theory* **68**, 155-174 (1992).
85. G. Borgefors and G. Santini di Baja, "Analyzing Nonconvex 2D and 3D Patterns", *Computer Vision and Image Understanding* **63**(1), 145-157 (1996).
86. Gabriella Sanniti di Baja, "Well-shaped, stable, and reversible skeletons from the (3,4)-distance transform", *Journal of Visual Communication and Image Representation* **5**(1), 107-115 (1994).
87. A.M. Vossepoel, "A Note on 'Distance Transformations in Digital Images'", *Computer Vision Graphics and Image Processing* **43**, 88-97 (1988).
88. G. Borgefors, "Another comment on 'A Note on 'Distance Transformation in Digital Images''", *CVGIP: Image Understanding* **54**(2), 301-306 (1991).

Indice de figuras

**Figura 1.1.** Dos imágenes adquiridas mediante el sensor Daedalus DS-1268 en las bandas espectrales (a) 0.45-0.52  $\mu\text{m}$  y (b) 2.08-2.35  $\mu\text{m}$ . (c) Resultado de su fusión, en el que se aprecian detalles presentes en las dos imágenes de entrada. (Imágenes cedidas por el Instituto Nacional de Técnica Aeroespacial, INTA). ..... 3

**Figura 2.1.** Comportamiento de los detectores de bordes y segmentos sobre imágenes de una misma escena en diferentes bandas espectrales. (a) Simulación de un par de imágenes IR-visible. (b) Bordes extraídos y (c) segmentos detectados. A-A', B-B' y C-C' indican segmentos correspondientes a los mismos bordes físicos de la escena pero que presentan configuraciones diferentes después del proceso de extracción..... 14

**Figura 3.1.** Diferencias en los conjuntos de segmentos extraídos de un par de imágenes de una misma escena tomadas en diferentes bandas espectrales. Fila superior: imagen IR, bordes extraídos y segmentos detectados; fila inferior: imagen visible, bordes y segmentos. .... 18

**Figura 3.2.** Una imagen IR (izquierda) y otra visible (derecha) de la misma escena, en la que puede apreciarse claramente el fenómeno de inversión de contraste. Las carreteras y algunos de los edificios, que aparecen claros sobre un fondo más oscuro en la imagen IR, se ven oscuros sobre fondo claro en la visible..... 19

**Figura 3.3.** (a) Imagen IR original, (b) bordes detectados, (c) imagen IR filtrada mediante relajación paramétrica y (d) bordes extraídos de la imagen filtrada con el mismo extractor de bordes utilizado para obtener (b)..... 21

**Figura 3.4.** (a) Imagen IR original y (b) bordes extraídos. (c) Resultado de la aplicación del algoritmo de reducción del bandeo a la imagen original y (d) bordes resultantes. .... 22

**Figura 3.5.** Comparación de la implementación usual del filtro de Canny (columna izquierda) y la implementación multirresolución propuesta (columna derecha), para anchuras de gaussiana,  $\sigma$ , (de arriba abajo) de 0.5, 5.0 y 10.0 píxeles. .... 24

**Figura 3.6.** (a) Imagen test, (b) bordes extraídos, (c) bordes adelgazados, (d) cadenas de píxeles conexos detectadas, (e) aproximación de las cadenas mediante listas de segmentos y (f) ajuste de *B-splines* cúbicos a las listas de segmentos. .... 26

**Figura 3.7.** Triángulos correspondientes (gris oscuro) en las imágenes de segmentos del par de imágenes de la figura 2.1. Los segmentos que definen los triángulos se han resaltado en blanco. La transformación que registra los triángulos también registra las imágenes completas. .... 29

**Figura 3.8.** (a) Un grupo de tres segmentos  $\{a, b, c\}$  definen un triángulo, aunque los segmentos no estén contenidos en los lados del mismo. (b) Una vez que los puntos de intersección de las rectas definidas por los segmentos han sido determinados, se etiquetan con subíndices crecientes según el sentido de las agujas del reloj..... 30



<b>Figura 3.9.</b> Correspondencia entre un triángulo fuente y uno destino. Según el orden de asignación de vértices se obtendrá cada una de las tres posibles transformaciones de registro. ....	30
<b>Figura 3.10.</b> Casamiento de dos parejas de segmentos, $\{s_1, s_2\}$ con $\{s_1', s_2'\}$ . Los vértices $A$ y $B$ de $s_1$ deben caer sobre la recta definida por $s_1'$ una vez que se aplique la transformación de registro $T$ . Análogamente, los vértices $C$ y $D$ de $s_2$ caerán sobre la recta definida por $s_2'$ . ....	33
<b>Figura 4.1.</b> (a) Imagen binaria, en la que se aprecian agregaciones de puntos a lo largo de dos rectas, y (b) su transformada de Hough. ....	38
<b>Figura 4.2.</b> Diferencia en los resultados de la acumulación (para un plano ángulo-escalado) en función de las precisiones de acumulación utilizadas: (a) acumulación con precisión superior a la requerida (tamaño de la matriz 2D, 108x144), que provoca dispersión, (b) precisión correcta (tamaño 36x29), (c) precisión ligeramente diferente a la necesaria (tamaño 32x32), que refuerza algunos máximos secundarios. ....	42
<b>Figura 4.3.</b> Mínima variación detectable en giro mediante la variación de la posición de un vértice equivalente a una longitud $dl$ a lo largo de uno de los ejes de coordenadas. ....	44
<b>Figura 4.4.</b> Mínima variación detectable en escalado mediante la variación de la posición de un vértice equivalente a una longitud $dl$ a lo largo de uno de los ejes de coordenadas. ....	45
<b>Figura 4.5.</b> Mínima variación detectable en sesgo mediante la variación de la posición de un vértice equivalente a una longitud $dl$ a lo largo de uno de los ejes de coordenadas. ....	45
<b>Figura 4.6.</b> Histogramas de las longitudes (de 0 a 100 píxeles) de los segmentos detectados en cuatro imágenes de prueba. ....	46
<b>Figura 4.7.</b> (a), (b) Dos ventanas de un par de imágenes IR-visible ya registradas. (c) Proyección de la región considerada del espacio de acumulación sobre el plano de traslaciones. (d) Proyección de la misma región una vez restringida la coordenada de giro. (e) Nueva proyección, después de restringir también el escalado. ....	50
<b>Figura 4.8.</b> Diferencias en los resultados de la acumulación según la función de peso utilizada. (a), (b) Par de imágenes de la misma banda espectral. (c), (d) Segmentos detectados. Planos de (e) escalado, (f) traslaciones y (g) ángulo/sesgo para la función de peso unidad. (h), (i), (j) Planos de acumulación empleando la longitud de superposición entre los segmentos que componen los triángulos casados como función de peso. ....	51
<b>Figura 4.9.</b> (a) Registro de la imagen 4.8a con la 4.8b utilizando la transformación de registro obtenida mediante la longitud de superposición como función de peso. (b) Fusión de 4.9a y 4.8b. ....	52
<b>Figura 4.10.</b> Registro de dos imágenes similares mediante acumulación con longitud de superposición como función de peso. (a) Imagen original, y (b) par de registro artificial, obtenido a partir de (a) mediante una rotación de $2^\circ$ y una traslación de $(-5, 2)$ píxeles. Planos de acumulación en (c) ángulo-sesgo y (d) traslaciones. (e) Imagen (b) registrada en función de los resultados de la acumulación. Las coordenadas del máximo de acumulación coinciden con las de la deformación aplicada (círculos). ....	53

- Figura 4.11.** Registro de dos imágenes de bandas espectrales alejadas (visible-IR) mediante acumulación con longitud de superposición como función de peso. (a) Imagen original, y (b) par de registro artificial, obtenido a partir de una imagen IR registrada con (a), mediante la aplicación de la misma transformación de desregistro utilizada en la figura 4.10. Planos de acumulación en (c) ángulo-sesgo y (d) traslaciones. (e) Imagen (b) registrada en función de los resultados de la acumulación. Los máximos de acumulación (círculos oscuros) no coinciden con los parámetros de la transformación aplicada (círculos claros). ..... 54
- Figura 4.12.** Registro de dos imágenes sintéticas mediante acumulación 4D en base a casamientos de parejas de segmentos, utilizando la superposición como función de peso. (a) Imagen sintética y (b) par de registro obtenido mediante una transformación afín de (a). Planos de corte en (c) escalado-ángulo y (d) traslaciones. (e) Resultado de la reconstrucción de la imagen 4.12b mediante la transformación correspondiente al máximo de acumulación detectado. .... 56
- Figura 5.1.** Longitud de superposición de dos segmentos. Para evaluarla se proyecta el segmento fuente transformado  $s_i$  sobre la recta definida por el segmento destino  $s_j$ . ..... 62
- Figura 5.2.** La distancia del segmento fuente transformado  $s_i$  al segmento destino  $s_j$  depende de su diferencia de orientación,  $d_\theta(s_i, s_j)$  y de la distancia desde el centro de  $s_i$  hasta la línea definida por  $s_j$ ,  $d_\rho(s_i, s_j)$ . ..... 64
- Figura 5.3.** Redundancias en la generación de triángulos. Los segmentos  $s_1$  y  $s_2$  forman el mismo triángulo al agruparse con  $s_3$  y  $s_4$ . Los segmentos colineales próximos se corresponden frecuentemente con el mismo borde fragmentado de la escena real. .... 67
- Figura 5.4.** Variación de  $Q(s_i, s_j)$  en función de la diferencia de orientación,  $d_\theta(s_i, s_j)$  y de la distancia desde el centro de  $s_i$  hasta la línea definida por  $s_j$ ,  $d_\rho(s_i, s_j)$ . A la derecha se muestra la misma gráfica en forma de curvas de nivel, trazadas para valores múltiplos de 0.1. .... 67
- Figura 5.5.** Esquema de la variación en la posición  $r$  del centro del segmento  $s_i$  a lo largo de la recta en la que está contenido, la cual forma un ángulo  $d_\theta(s_i, s_j)$  con la recta que contiene al segmento destino  $s_j$ . ..... 68
- Figura 5.6.** Valor de la calidad de casamiento entre los segmentos fuente y destino cuando la posición relativa de los mismos varía según el esquema de la figura 5.5. A la derecha, el mismo gráfico en forma de curvas de nivel, cada 0.1. .... 68
- Figura 5.7.** Esquema de la variación en la orientación  $\phi$  del segmento  $s_i$  respecto a la orientación de la recta en la que está contenido su centro, la cual forma un ángulo  $\theta$  con la recta que contiene al segmento destino  $s_j$ . El centro del segmento  $s_i$  se mantiene a una distancia fija,  $r$ , del punto de corte de las rectas, como en el diagrama de la figura 5.5. .... 69
- Figura 5.8.** Valor de la calidad de casamiento entre los segmentos fuente y destino cuando la posición relativa de los mismos varía según el esquema de la figura 5.7, para una distancia  $r=20$  píxeles. A la derecha, el mismo gráfico en forma de curvas de nivel, cada 0.1. .... 69
- Figura 5.9.** Registro de una imagen IR con una visible de la misma escena. (a) Imagen IR, (b) imagen visible, (c) segmentos detectados en la imagen IR y (d) en la visible. (e) Segmentos de

la imagen IR casados con la transformación afín óptima y (f) imagen IR registrada correspondiente. (g) Fusión de la imagen IR registrada con la imagen visible.....	72
<b>Figura 6.1.</b> (a) Registro de la imagen 4.8a con la 4.8b mediante un polinomio de tercer orden, obtenido como refinamiento de la transformación afín determinada por el máximo de acumulación mostrado en las imágenes 4.8h, i, j. (b) Resultado de la fusión con la imagen 4.8b. ....	78
<b>Figura 6.2.</b> (a) Registro de la imagen IR 5.9a con una imagen visible de la misma escena (5.9b) mediante el método de casamiento de conjuntos de segmentos, y (b) fusión resultante. (c) Imagen IR registrada mediante la corrección de tercer orden de la transformación afín óptima y (d) fusión con la imagen visible. ....	79
<b>Figura 7.1.</b> (a) Fusión de dos imágenes de bandas espectrales próximas ligeramente desregistradas, en la que se aprecia claramente la aparición de bordes dobles, debida a la falta de correspondencia espacial precisa entre las dos imágenes. (b) Fusión de las imágenes una vez que han sido correctamente registradas. ....	83
<b>Figura 7.2.</b> Componente de Gabor de alta frecuencia, $G_I$ , obtenida mediante el filtro $M_{7 \times 7}$ . (a) Imagen original, (b) componente $G_I$ (con signo) y (c) módulo de la componente de Gabor. ....	85
<b>Figura 7.3.</b> Recubrimiento del espectro de frecuencias para la componente de Gabor de frecuencia más alta. En la fila superior se muestra el recubrimiento combinado de las funciones de Gabor en frecuencia para las cuatro orientaciones necesarias. En la fila de abajo puede verse el recubrimiento correspondiente a la máscara integrada $M_{7 \times 7}$ . ....	86
<b>Figura 7.4.</b> Borde ideal, utilizado para determinar los límites de aplicabilidad del criterio de registro. (a) Gradiente orientado a $0^\circ$ y (b) a $45^\circ$ . ....	91
<b>Figura 7.5.</b> Incrementos de energía $\overline{\Delta E}(r)$ para el borde ideal. (a) Curva para el borde orientado a $0^\circ$ . (b) Curva para el borde orientado a $45^\circ$ . ....	91
<b>Figura 7.6.</b> (a), (b) Par de imágenes sintéticas ( $512 \times 512$ píxeles) generadas a partir del diagrama de Voronoi de una distribución de puntos aleatoria. (c) Resultado de su fusión. ....	92
<b>Figura 7.7.</b> (a), (b) Regiones centrales (de $256 \times 256$ píxeles) de las imágenes sintéticas de la figura 7.6, utilizadas para evitar efectos de borde. (c) Fusión de (a) y (b). ....	93
<b>Figura 7.8.</b> Mapa de incrementos de energía asociados a la traslación de la imagen sintética 7.7a sobre la 7.7b. ....	94
<b>Figura 7.9.</b> Incremento de energía $\overline{\Delta E}(r)$ en función de la distancia para el par sintético de la figura 7.7, obtenido a partir del mapa de incrementos de la figura 7.8. ....	94
<b>Figura 7.10.</b> Fusión del par de imágenes sintético después de aplicar las deformaciones de prueba: (a) traslación de (2, -2) píxeles, (b) rotación de $2^\circ$ , (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%. ....	95
<b>Figura 7.11.</b> Bandas (a) de 2.08 a 2.35 $\mu\text{m}$ y (b) de 0.45 a 0.52 $\mu\text{m}$ de una imagen Daedalus de Navalcarnero. (c) El resultado de su fusión. ....	97

<b>Figura 7.12.</b> (a), (b) Regiones centrales (de 256x256 píxeles) de las imágenes Daedalus de la figura 7.11, utilizadas para evitar efectos de borde. (c) y su fusión. ....	98
<b>Figura 7.13.</b> Mapa de incrementos de energía del par de imágenes Daedalus de la figura 7.12. ....	99
<b>Figura 7.14.</b> Incremento de energía $\Delta E(r)$ en función de la distancia para las imágenes de la figura 7.12, obtenido a partir del mapa de incrementos de la figura 7.13. ....	99
<b>Figura 7.15.</b> Fusión del par de imágenes Daedalus después de aplicar las deformaciones de prueba: (a) traslación de (2, -2) píxeles, (b) rotación de 2°, (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%. ....	100
<b>Figura 7.16.</b> Dos formas de distribuir puntos en una imagen para la evaluación del valor de desplazamiento medio equivalente. (a) Rejilla cuadrada y (b) malla distribuida radialmente, con el mismo número de puntos de muestra en cada anillo. ....	102
<b>Figura 7.17.</b> Valor del peso $\chi$ en función del factor de forma $q$ y de la diferencia de orientación $\alpha$ entre un vector de desplazamiento y la dirección del eje principal de la elipse de confianza. ....	105
<b>Figura 7.18.</b> Campos vectoriales correspondientes a las funciones de deformaciones de prueba aplicadas a la imagen fuente sintética 7.5a. (a) traslación de (2, -2) píxeles, (b) rotación de 2°, (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%. ....	107
<b>Figura 7.19.</b> Campos vectoriales correspondientes a las funciones de deformación aplicadas a la imagen fuente Daedalus 7.10a. (a) traslación de (2, -2) píxeles, (b) rotación de 2°, (c) sesgo de 0.03, (d) escalado del 103%, (e) distorsión en corsé del 3%, (f) distorsión en barrilete del 3%. ....	108
<b>Figura 7.20.</b> Imágenes IR registradas y resultados de su fusión con la imagen visible. (a) y (b) Par de registro IR-visible. (c) y (d) Registro y fusión para el modelo afín de transformación. (e) y (f) Registro y fusión para el modelo de tercer orden. (g) y (h) Resultados para la imagen de referencia óptima, generada mediante registro manual. ....	111
<b>Figura 7.21.</b> Componentes de Gabor de alta frecuencia, $G_I$ . (a) Región de interés. (b) Imagen visible. (c) IR registrada con el modelo afín. (d) IR registrada con el modelo de tercer orden. (e) IR de referencia. (f), (g) y (h) Componentes de Gabor para las fusiones de las imágenes IR (c), (d) y (e) con la imagen visible. ....	112
<b>Figura 7.22.</b> Mapa de incrementos de energía generado para la zona de interés (7.21a), a partir de la imagen IR de referencia (7.20g) y la imagen visible (7.20b). El punto de mínima energía coincide con el de traslación nula (0, 0). ....	112
<b>Figura 7.23.</b> Curva de incrementos de energía $\Delta E(r)$ en función de la distancia al punto mínimo, obtenida a partir del mapa de incrementos de la figura 7.22. ....	113
<b>Figura 7.24.</b> Campos vectoriales de desplazamientos locales, con respecto a la imagen de referencia óptima (7.20g), para las dos imágenes infrarrojas registradas de forma automática. (a) Imagen IR registrada mediante el modelo afín. (b) Imagen IR registrada mediante el modelo de tercer orden. ....	113

<b>Figura 8.1.</b> Dos imágenes, IR (a) y visible (b) de una misma escena. (c) Registro automático de (a) con (b) basado en calidad de casamientos con refinamiento de tercer orden, y (d) resultado de la fusión de (c) y (b). (e) Registro manual y (f) resultado de su fusión con (b). .....	116
<b>Figura 8.2.</b> Mapa y curva de incrementos de energía para la evaluación del registro del par IR-visible de la imagen 8.1.....	117
<b>Figura 8.3.</b> Campo vectorial de desplazamientos locales para la imagen 8.1c, registrada automáticamente, respecto de la registrada manualmente, 8.1e.....	117
<b>Figura 8.4.</b> Registro de dos imágenes aéreas en bandas IR (a) y visible (b). (c) Registro automático de tercer orden y (d) fusión resultante. (e) Registro manual y (f) fusión. ....	118
<b>Figura 8.5.</b> Mapa y curva de incrementos de energía para la evaluación del registro del par IR-visible de la figura 8.4. ....	119
<b>Figura 8.6.</b> Campo vectorial de desplazamientos locales para la imagen 8.4c, registrada automáticamente, respecto de la registrada óptima, correspondiente a la traslación de (-1, 0) píxeles de la registrada manualmente, 8.4e. ....	120
<b>Figura 8.7.</b> Regiones de medida para refinar la posición de un segmento, de longitud $l$ , sobre la imagen SAR. Las regiones de medida son de forma rectangular, de altura $l$ y anchura $a$ . y las poblaciones de niveles de gris contenidas en ellas se modelan mediante distribuciones gaussianas, $(\mu_1, \sigma_1)$ y $(\mu_2, \sigma_2)$ . ....	126
<b>Figura 8.8.</b> Variación del valor del discriminante de Fisher, $D$ , según se modifica la posición del segmento perpendicularmente a la frontera de separación de dos regiones de medias y varianzas diferentes. El máximo se produce cuando la posición del segmento coincide con la frontera de separación de las regiones. En este ejemplo $\mu_1 = \sigma_1^2 = 100$ , $\mu_2 = \sigma_2^2 = 200$ , $a = 4$ píxeles y $l = 12$ píxeles. ....	128
<b>Figura 8.9.</b> Ajuste fino sobre la imagen SAR de un segmento transformado mediante la función de registro aproximada. (a) Segmento detectado en la imagen visible. (b) Segmento transformado sobre la imagen SAR con la función de transformación aproximada. (c) Las dos regiones de medida a los lados del segmento transformado, en la posición en la que el nivel de discriminación entre sus poblaciones es máximo.....	129
<b>Figura 8.10.</b> Registro visible-SAR mediante el método propuesto. (a) Región de interés en la imagen visible. (b) Región correspondiente de la imagen SAR, determinada mediante la transformación de registro aproximada. (c) Segmentos detectados en la imagen visible. (d) Segmentos transformados sobre la imagen SAR utilizando la transformación aproximada. (e) Región SAR re-extraída de la imagen completa utilizando la transformación corregida. Como puede verse, los segmentos transformados coinciden ahora de forma precisa con los bordes de la imagen SAR.....	130
<b>Figura 8.11.</b> Registro (a) automático y (b) manual de la imagen SAR 8.10a con la visible 8.10b, utilizados para la determinación de la calidad en función del criterio de Gabor.....	131

<b>Figura 8.12.</b> Mapa de incrementos de energía de la imagen SAR registrada automáticamente con la visible. Debajo, curva de desplazamientos equivalentes.....	132
<b>Figura 8.13.</b> Aplicación del registro al reconocimiento de formas. (a) Silueta del objetivo y (b) imagen de búsqueda. (c) Silueta registrada y (d) su fusión con (b). ....	133
<b>Figura 8.14.</b> Aplicación del registro a la generación de mosaicos de imágenes. (a) y (b) Dos fotografías de la parte norte de Vitoria. (c) Registro de (a) con (b) y (d) fusión de (c) con (b). ....	134
<b>Figura 8.15.</b> Aplicación del registro a la fusión de mapas y fotografías. (a) Mapa y (b) fotografía aérea de una misma zona. (c) Mapa umbralizado, utilizado como imagen de bordes. (d) Mapa registrado con la fotografía y (e) resultado de la fusión de ambos.....	135
<b>Figura B.1.</b> Aplicación del filtro desarrollado. (a) Imagen IR con ruido de barrido y (b) bordes extraídos mediante Canny ( $\sigma = 3.0$ píxeles). (c) Resultado de la aplicación del filtro y (d) bordes resultantes. ....	145
<b>Figura C.1.</b> Aplicación de los diferentes modelos de crecimiento a un píxel aislado situado en la esquina inferior izquierda de las imágenes, y el arco circular ideal para una circunferencia de radio 255 píxeles. (a) Crecimiento en 4-v, (b) 8-v, (c) crecimiento octogonal (iteración de 4-v, 8-v), (d) crecimiento octogonal regular y (e) crecimiento hexadecagonal propuesto. Sólo se muestran los primeros cuadrantes. ....	148
<b>Figura C.2.</b> Tipos de vecindades para los píxeles del borde del octógono. (a), (b) Máscaras para los píxeles lado. (c), (d) Máscaras para los píxeles vértice. ....	151
<b>Figura C.3.</b> (a) Resultado de la eliminación de vértices (2 iteraciones). Los píxeles eliminados se muestran en color gris oscuro. (b) Resultado de la inhibición del crecimiento de los píxeles vértice (2 iteraciones). Los píxeles que no crecieron en la iteración correspondiente se muestran en gris oscuro. El resultado final es equivalente al de la figura C.3a. ....	151
<b>Figura C.4.</b> Truncación de un octógono regular circunscrito a una circunferencia de radio $n$ mediante una recta de pendiente -2. ....	152
<b>Figura C.5.</b> Comparación de los diferentes modelos de dilatación 3x3 analizados, después de 24 iteraciones. (a) Imagen binaria con algunas formas aisladas. (b) Dilatación mediante 4-v. (c) Dilatación por 8-v. (d) Dilatación por crecimiento octogonal (4,8). (e) Simulación de dilatación mediante distancia Chamfer 3/4. (f) Dilatación por crecimiento octogonal regular, y (g) modelo de crecimiento hexadecagonal propuesto. ....	155
<b>Figura C.6.</b> Diagrama de Voronoi de un conjunto de figuras binarias arbitrarias, obtenido mediante el modelo de crecimiento hexadecagonal propuesto. ....	156
<b>Figura C.7.</b> Mapas de distancia para los segmentos de la figura C.5a. (a) Distancia de Chamfer 3/4, (b) distancia de Chamfer 5/7/11, (c) distancia hexadecagonal utilizando el modelo propuesto y (d) discretización de la distancia euclídea. Para apreciar mejor el patrón de crecimiento se muestran las distancias módulo 32. Las imágenes son de 256x256 píxeles y los errores máximos para cada uno de los modelos son aquí de 12.39 píxeles para el Chamfer 3/4, 3.44 para el Chamfer 5/7/11 y 2.16 para el crecimiento hexadecagonal propuesto. ....	156



Indice de tablas.

**Tabla 7.1.** Máscara integrada,  $M_{7 \times 7}$ , utilizada para el filtrado de Gabor. .... 86

**Tabla 7.2.** Desplazamientos equivalentes (en píxeles) para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes sintéticas. .... 96

**Tabla 7.3.** Desplazamientos equivalentes (en píxeles) para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes Daedalus. .... 98

**Tabla 7.4.** Módulos medios y desplazamientos medios equivalentes (ambos en píxeles), calculados en función del campo vectorial de desregistro, para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes sintéticas. .... 106

**Tabla 7.5.** Módulos medios y desplazamientos medios equivalentes (ambos en píxeles), calculados en función del campo vectorial de desregistro, para cada una de las deformaciones de prueba aplicadas a la pareja de imágenes Daedalus. .... 106

**Tabla 7.6.** Desplazamientos medios equivalentes (en píxeles), para los dos modelos de transformación en base a los criterios de evaluación desarrollados. .... 110

**Tabla 8.1.** Desplazamiento medio equivalente mediante el criterio de Gabor, media del módulo de los vectores del campo de desplazamientos locales y desplazamiento equivalente en función del campo vectorial, para el registro de las imágenes de la figura 8.1. Medidas en píxeles. .. 116

**Tabla 8.2.** Desplazamientos equivalente para los registros manual y automático de la imagen IR 8.4a con la visible 8.4b, en función del criterio de Gabor. .... 119

**Tabla 8.3.** Desplazamientos equivalentes en función del campo vectorial de desplazamientos locales para la imagen 8.4c y la imagen registrada óptima, correspondiente a la traslación de (-1,0) píxeles de la imagen 8.4e. Medidas en píxeles. .... 120

**Tabla 8.4.** Desplazamientos equivalentes para los registros manual y automático de la imagen SAR 8.10a con la visible 8.10b, en función del criterio de Gabor. .... 131

**Tabla C.1.** Cocientes entre el área y el cuadrado del radio, errores de área y errores de distancia para cada uno de los modelos de crecimiento analizados. .... 154





**Notas**

